THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

~~FEBRUARY~~ 1968

VOL. 38, ~~PART 1~~

Price One Pound

10.3.81
J 754

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
J. D. NISBET
Assistant Editor :
R. D. DUTCH
Associate Editors :
P. E. VERNON and DORIS M. LEE
Advised by
SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. 38, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 11, New Fetter Lane, London, E.C.4. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to Professor J. D. Nisbet, King's College, Aberdeen. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to Professor J. D. Nisbet, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Sweet and Maxwell, Spon Booksellers, Ltd., 11, New Fetter Lane, London, E.C.4.

Published by
METHUEN AND CO., LTD., 11, NEW FETTER LANE, LONDON, E.C.4,
for the British Journal of Educational Psychology, Ltd.
Registered Offices : 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

VOLUME 38 FEBRUARY 1968 PART 1

185

SYMPOSIUM: RECENT RESEARCH ON MALADJUSTMENT

SUMMARY. A symposium on "Recent Research on Maladjustment" was convened by Maurice Chazan and chaired by Professor Jack Tizard at the Annual Conference of the Education Section of the British Psychological Society in September, 1967. Summaries of the four papers in the symposium are presented here, together with a post-script which Professor Tizard has added.

THE BEHAVIOUR AND ADJUSTMENT OF SEVEN-YEAR-OLD CHILDREN: SOME RESULTS FROM THE NATIONAL CHILD DEVELOPMENT STUDY (1958 COHORT)

BY RONALD DAVIE

(National Child Development Study)

THE National Child Development Study is a longitudinal study of a national cohort of some 16,000 children born in 1958. The population of children involved is one week's births in England, Scotland and Wales and they were the subjects of an extensive and detailed survey at the time of their birth (Butler and Bonham, 1963). The cohort was retraced and information collected in 1965 on many aspects of their education, behaviour, health and environment. At that time the children were in the final term of their infant schooling; and the results presented in this abstract relate to a sub-sample of approximately 11,000 children in England. The full results of the first report have been published as a paperback (Pringle, Butler and Davie, 1967).

A number of assessments were made of the children's behaviour and adjustment: teachers completed a Bristol Social-Adjustment Guide (Stott, 1963) for each of the children and they were also asked to rate how quickly the children settled down on first starting school; mothers were posed a similar question about the settling down period, and the occurrence of specific aspects of current behaviour was also enquired. The results are not easy to summarize adequately but three general findings are of particular interest.

First, a comparison of the sexes showed a much higher incidence of behavioural deviance amongst the boys. For example, twice as many boys as girls had a total B.S.A.G. score of 20 or over (termed by Stott 'maladjusted'); the proportion of boys still unsettled at school three months after starting was 7 per cent., compared with a figure of 5 per cent. for the girls (as reported by teachers); and the comparable reports from mothers showed a similar but slightly wider gap.

In general, boys were more often reported to show current 'difficulties' at home. However, it is interesting to note that with one or two exceptions, the sex differences, taken singly, although statistically significant, were not large enough to be clinically useful in terms of any differential significance of presenting symptoms for boys and girls. Nevertheless, the overall picture indicates fairly clear sex patterns: the boys showed more aggression (temper tantrums, fighting, disobedience, destructiveness), more restless behaviour (difficulty in settling to tasks, fidgeting) and irritability; the girls displayed more overt anxiety (worrying, tearfulness), feeding difficulties (faddiness, poor appetite) and 'oral activity' (nail-biting, thumb and finger-sucking).

Secondly, there were clear social class differences in relation to social-adjustment scores. For example, the proportion of ' stable ' boys (B.S.A.G. score 0–9) in Social Classes 1 and 2 together was 69 per cent., whilst in Social Class 5 it was 43 per cent. The corresponding percentages for girls were 79 and 62. However, whilst the fall through the Social Classes in boys showed no significant departure from linearity, there was a highly significant departure from linear trend in the girls' results. It would appear, then, that there is a differential relationship between socio-economic status and social-adjustment for boys and girls.

Finally, there was evidence—tentative as yet—of a relationship between the period of gestation and subsequent adjustment. No statistical analyses have so far been carried out but first results suggest that, after allowance is made for Social Class, those children born at term (39–41 weeks) are the most stable, whilst those born both pre-term and post-term show more evidence of poor adjustment. A similar relationship has been shown in this study between gestational maturity and tested reading performance. Further analyses are needed to investigate whether these relationships remain when allowance is made for other social and obstetric factors. However, Barker and Edwards (1967) in an investigation of 11-year-olds, using siblings as controls, produced quite strong evidence that the measured intelligence of children born at term is higher than that of children born both prematurely and postmaturely.

REFERENCES.

BARKER, D. J. P., and EDWARDS, J. H. (1967). Obstetric complications and school performance. *Brit. Med. J.*, 3, 695-699.

BUTLER, N. R., and BONHAM, D. G. (1963). *Perinatal Mortality*. Edinburgh : Livingstone.

PRINGLE, M. L. K., BUTLER, N. R., and DAVIE, R. (1967). 11,000 *Seven-year-olds*. London : Longmans Green.

STOTT, D. H. (1963). *The Social Adjustment of Children : Manual to the Bristol Social-Adjustment Guides*. London : Univ. of London Press.

ADJUSTMENT AND EDUCATIONAL PROGRESS*

BY J. W. B. DOUGLAS AND J. M. ROSS
(*National Survey of Health and Development*)

THE title of this paper, "*Adjustment* and Educational Progress," was purposely chosen, for we are not primarily concerned with the study of the small group of boys and girls who are receiving or ought to be receiving some form of psychiatric help. The focus of our attention has been the prevalence of signs of abnormal behaviour in a representative population of school children and the study of the links between general adjustment, as so assessed, and educational progress and attainment.

---

* This short summary is based on a chapter in the book, *All Our Future*, by J. W. B. Douglas, J. M. Ross and H. R. Simpson, to be published this year.

The National Survey of Health and Development has detailed records of the home circumstances, growth, illnesses and school records of some 4,000 boys and girls who have been closely followed from their birth in 1946 to the present day. Our information on adjustment comes from three main sources, the teachers, the mothers, and the boys and girls themselves. The first of these assessments comes from the secondary school teachers' reports on out of class behaviour in two separate years, the second from detailed accounts of symptoms and habits given to us at 6, 11 and 15 years, and the third from a self-rating inventory based on the Pintner, completed at 13 years, designed to measure levels of neuroticism, intro- or extraversion and aggression. The three measures of adjustment, i.e., teachers' assessments, reported symptoms and level of neuroticism, pick out children who are in many other ways unusual, for example, the truants, delinquents, those who had a number of accidents, and those who from independent reports by their mothers and primary school teachers have shown various behaviour disturbances.

We have classified the boys and girls by their out of class behaviour and symptom counts at 15 years into three groups: the 'well adjusted,' 27 per cent. of the total; the 'less well adjusted' showing some signs of abnormal behaviour, 45 per cent.; and the 'least well adjusted,' 28 per cent. We have further divided each of these three categories into two by the neuroticism scores made on the self-rating inventory . The school progress of these six groups is very different. The average test scores of the 'well adjusted non-neurotic' pupils is four-fifths S.D. above that of the 'least well adjusted neurotic.'

All the four tests used at 15 years show a similar pattern except that in the most disturbed group (6 per cent. of the sample), the verbal and attainment scores are considerably below the non-verbal test scores. The average test score differences at 8 and 11 years between the two extreme adjustment groups are 73 per cent. and 86 per cent. as large respectively as the 15-year one. In fact, it is only these two extreme groups that show any relative change in their test score results over time—the 'well adjusted non-neurotic' improving and the 'least well adjusted neurotic' deteriorating by roughly similar amounts. The scores of the intermediate groups, however, remain at similar levels in each test battery.

The conclusion to be drawn from this study of the relation between behaviour and school progress is that it is not only the grossly disturbed pupils who are affected. There is a continuous gradient in ability and attainment, running from those who have no adverse teachers' comments, no symptom reported and low neuroticism scores to those who are picked out by all three measures. The greater the number of adverse items reported, the lower, on the whole, is school performance. Although there is evidence of a deterioration in relative performance during the school years, this is confined to a small group of very maladjusted children; for the majority, the increase in symptoms or signs of disturbed behaviour is related to lower performance at *each* age tested. This would suggest that the basis of their educational difficulties lies in the pre- or early school years.

REFERENCE.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London: MacGibbon and Kee.

TABLE 1

AGGREGATE TEST SCORES* RELATED TO A COMBINATION OF NEUROTICISM, BEHAVIOUR RATINGS AND SYMPTOMS AT 15.
(Based on population estimates.)

	Neuroticism Scores					
	High			Other		
	Behaviour Ratings and Symptoms			Behaviour Ratings and Symptoms		
	High	Average	Low	High	Average	Low
Test score at age : 8	46·2	49·4	51·2	48·6	50·6	52·1
11	45·9	49·2	51·8	48·4	50·8	52·9
15	45·2	49·9	51·6	48·1	51·1	53·3
Percentage of sample (percentaged across) ..	6	14	13	18	33	16

*The aggregate test scores were obtained by summing the 4 standardized scores at each age and re-standardizing these on the population that sat all three test batteries to a mean of 50 and a S.D. of 9.5

TABLE 2

TEST SCORES* AT 15 RELATED TO A COMBINATION OF NEUROTICISM, BEHAVIOUR RATINGS AND SYMPTOMS AT 15.

(Based on population estimates.)

Test Scores at age 15.	Neuroticism Scores					
	High			Other		
	Behaviour Ratings and Symptoms			Behaviour Ratings and Symptoms		
	High	Average	Low	High	Average	Low
Non-verbal	47·8	50·8	51·2	49·1	51·5	53·1
Verbal	45·7	50·3	51·9	48·4	51·1	52·6
Reading	46·3	49·7	52·2	48·3	50·9	53·7
Mathematics	45·6	50·6	51·7	49·3	51·8	53·3

* Each test was standardized on the population taking the test to a mean of 50 and a S.D. of 10.

INCONSEQUENTIAL BEHAVIOUR IN SCHOOL CHILDREN

BY MAURICE CHAZAN

(Department of Education, University College of Swansea)

The purpose of this contribution to the symposium on maladjustment is to stimulate some discussion on what Stott (1966, a, b) has termed the ' inconsequential ' syndrome in school children. This syndrome is an extension of the concept of ' restlessness ' used in the analysis of the Bristol Social Adjustment Guide (Child in School), and its thirteen core symptoms are listed in the Appendix. Children who show inconsequential behaviour display a failure in temporal integration (Hearnshaw, 1956) ; they characteristically take only the immediate situation into account, and lack control in response to petty frustrations. They are highly distractible and do not maintain interest in anything for long.

Such children present a serious problem in the classroom and further investigation, and possibly refinement, of the syndrome would appear to be of value. As a preliminary study, the incidence of inconsequential behaviour was looked at in three samples of children (907 in all), in South Wales, for whom completed Bristol Guides were available :

(i) 470 pupils (225 boys, 245 girls) in ordinary junior and secondary school classes.

(ii) 99 pupils (64 boys, 35 girls) in special *classes* for E.S.N. children attached to junior schools.

(iii) 169 pupils (110 boys, 59 girls) in special *schools* for the E.S.N., and a control group of 169 pupils in ordinary schools.

The criterion adopted for ' inconsequential ' behaviour was the underlining by the teacher of at least *six* of the thirteen inconsequential items listed in the Appendix. In the sample of children attending ordinary schools, the incidence of inconsequential behaviour was low, only 10 out of 470 pupils (2·1 per cent.) scoring six or more inconsequential points. It was higher (5·9 per cent.) in the sample of E.S.N. children in special *schools*, there being significantly more cases of inconsequential behaviour in this sample than in the control group of children in ordinary schools (1·8 per cent.), and particularly high (17 per cent.) in the sample of E.S.N. pupils (aged 8 to 12 years) in special *classes*. No statistically significant sex or age differences were found, though there was a tendency for boys to show a greater amount of inconsequential behaviour, and for its incidence to be higher in junior than in secondary school children.

Children scoring high on the inconsequential syndrome tended to have high total maladjustment scores. Inconsequential behaviour was, in nearly every case found in the three samples, associated with some degree of educational failure. Only one of the ten inconsequential pupils in the E.S.N. school sample had made satisfactory scholastic progress at the special school, and even in the ordinary school sample four out of the six inconsequential boys were very backward, though all four girls had reasonable attainments.

The concept of inconsequential behaviour appears to be a useful one, particularly in describing children who are restless and lacking in consistent purposefulness, but who behave in a less extreme way than the children described as ' hyperkinetic ' in the psychiatric literature. Rutter (1965), in fact, considers that the term ' hyperkinetic ' properly describes a syndrome characterized by *severe* and *disorganized* overactivity together with impulsiveness, a very short attention span and often extreme distractibility. The inconsequential syndrome, as suggested by Stott (1966, b) seems, however, to need some modification. Analysis of the completed Guides of the forty inconsequential children suggests that items R1 and XC15 (see Appendix) might be excluded because of their low frequency in these records, and that items XC11 and D1 might be omitted because these items are frequently marked by teachers in the case of children who cannot be regarded as showing inconsequential behaviour. These items might be replaced by the following items, on the grounds that they occur frequently in inconsequential records and that they fit meaningfully into the syndrome :

D 8 — Flies into a temper if provoked.
XC 3 — Inclined to fool around (team games).
XC 10 — Shows off (pulls silly faces, mimics, clowns).
HA 14 — Resentful muttering or expression at times.
HC 4 — Squabbles, makes insulting remarks.

The modified syndrome of inconsequential behaviour would thus consist of these fourteen items from the Bristol Social Adjustment Guide : R2, 5, 6, 9, 10, 11 ; XA10 ; XC3, 10, 14 ; D8, 13 ; HC4 ; HA14.

It may be asked why children in special *classes* for the E.S.N. show so much more inconsequential behaviour than those in special *schools*. A possible reason is that the special school is usually able to provide a much more varied programme of activities than the special class, which is often restricted in space and materials.

It should be stressed that it is not suggested here that children showing some degree of inconsequential behaviour represent a clear-cut type ; as stated above, the children in the sample showed a variety of other symptoms of maladjustment as well as those included in the inconsequential syndrome. Nor do we know very much about the causes of such behaviour, though Stott (1966), while recognizing that more work needs to be done on both the nature and causes of the syndrome, considers that inconsequential behaviour is mainly due to some congenital neural impairment. Nevertheless, it is particularly in the case of such children, who resist normal methods of discipline and correction, that the teacher needs help. Neither punishment nor reward seems to have much effect on their behaviour, and it is not easy for teachers to maintain a sympathetic and tolerant attitude towards children who provoke them continually. The situation is especially difficult when there are several inconsequential children in the same class, as is not unlikely in a special school or class for E.S.N. children. It is important, therefore, that further attention should be given to possible methods of approaching the problem of inconsequential behaviour.

References.

Hearnshaw, L. S. (1956). Temporal integration and behaviour. *Bulletin of the Brit. Psychol. Soc.*, 30, 1-20.

Rutter, M. (1965). Classification and categorization in child psychiatry. *J. Child Psychol. Psychiat.*, 6, 71-83.

Stott, D. H. (1966a). *The Social Adjustment of Children : Manual to the Bristol Social Adjustment Guides* (3rd edition). Univ. London Press.

Stott, D. H. (1966b). *Studies of Troublesome Children*. London : Tavistock Publications.

APPENDIX.

CORE SYMPTOMS OF INCONSEQUENTIAL SYNDROME (Stott, 1966, a).

Coding on Bristol Guide	*Item.*
R (*Restlessness*)	
1	Gets very dirty during day.
2	Starts off others in scrapping and rough play.
5	Careless, untidy ; often loses or forgets books, pens.
6	Rough-and-ready, slapdash (in manual work).
9	Cannot attend or concentrate for long.
10	Does not know what to do with himself, can never stick at anything long.
11	Too restless to remember for long (after correction).
XA (*Anxiety about adult interest and affection*).	
10	Constantly needs petty correction.
XC (*Anxiety for approval of and acceptance by other children*).	
11	Misbehaves when teacher is out of room.
14	Foolish pranks when with a gang.
15	Follower in mischief.
D (*Depression*).	
1	Sometimes eager, sometimes doesn't bother (answering questions).
13	Slumps, lolls about.

EDUCATIONAL ASPECTS OF CHILDHOOD MALADJUSTMENT : SOME EPIDEMIOLOGICAL FINDINGS

BY WILLIAM YULE AND MICHAEL RUTTER

(*Institute of Education and Institute of Psychiatry, University of London*)

THE nature and extent of overlap between physical, social, psychiatric, cognitive and educational handicaps were among the prime problems examined in the Isle of Wight survey of some 2,200 9 to 11-year-old children (Rutter, *et al.*, 1966). In this paper, however, we will only be concerned with one part of this problem—that is the educational and cognitive correlates of maladjustment.

With each handicap studied, a similar two-stage approach was adopted. Firstly, children who might suffer from any of the relevant conditions were chosen by means of a series of screening procedures. Secondly, these possibly handicapped children were studied individually by physicians, psychiatrists, psychologists and social scientists, and reports were obtained from parents, schools, hospitals, etc. The final selection of handicapped children was then made on the basis of these intensive examinations.

Thus, behavioural questionnaires of demonstrated reliability and validity (one completed by teachers and one by parents) were used as the main screening techniques for maladjustment (Rutter, 1967 ; Rutter and Graham, 1966).

In addition, children who had been before the courts, were in long-term care, or were currently attending a psychiatrist (all small groups) were also chosen for individual study, making a total group of 284 possibly maladjusted children. These 284 children were then seen individually by psychiatrists and psychologists, the parents were interviewed and further information was obtained from schools. On this basis a final judgment as to maladjustment was made by psychiatrists for each child (Rutter and Graham, 1967). One-hundred-and-twenty-six were considered to have some clinically important psychiatric disorder—that is to say, they were maladjusted. This represents a minimum prevalence rate of 5·7 per cent. (children with minor emotional problems, uncomplicated mental subnormality or monosymptomatic disorders like nocturnal enuresis were not represented in this rate.)

The two largest diagnostic groups among the maladjusted children were neurotic disorders and antisocial disorders (Rutter, 1965). There was a marked difference in sex distribution in these groups; while there were slightly fewer boys than girls among the children with neurotic disorders (17 : 26), there was a very marked excess of boys among those with antisocial disorders (34: 9). Overall, there was a higher prevalence of maladjustment in boys (64·5 per cent. of the 126 were boys).

Scores on a short form of the WISC showed the maladjusted group to be of lower intelligence than a normal control group (the difference between the means was half a standard deviation: $p < \cdot 001$). On Neale's (1958) reading test the maladjusted children were on average 19 months retarded in accuracy and 17 months retarded in comprehension compared with their chronological age. This educational retardation was *not* simply a reflection of their somewhat lower intelligence since, when intelligence was partialled out, nearly a quarter (23·8 per cent.) of the psychiatric group were found to be at least twenty-eight months retarded in their reading in relation to their I.Q. and chronological age in comparison with only 5·4 per cent. of the control group (Yule, 1967).

When non-verbal I.Q. and reading attainment (assessed on N.F.E.R. group tests) were examined on the total population* in relation to individual items of deviant, but not necessarily pathological behaviour as reported by parents and teachers, equally strong and highly significant associations were found. The trends were linear, showing that the associations applied across the whole range of intelligence and attainment.

However, there were important differences between diagnostic groups (among the 126 with significant psychiatric disorder) with regard to reading attainment. Whereas over a third of the antisocial children were severely retarded (twenty-eight months or more behind their predicted reading age), reading retardation among neurotic children was little more frequent than in the general population. This does not merely reflect the difference in sex composition between the groups; of the thirty-two antisocial boys tested, thirteen were severely retarded in reading, but none of the seventeen neurotic boys were so retarded.

Thus, there appears to be a strong association between poor reading attainment and antisocial disorder but little association between reading attainment and neurosis. The study was able to provide only circumstantial evidence on the nature of the association, but the findings did suggest that it was rather unlikely that the psychiatric disorders as such led to reading failure. Maladjusted children who were backward in reading shared much the same developmental delays in language and perception and much the same social characteristic as the poor readers who were *not* maladjusted, while they did not share many of the family characteristics of the maladjusted children who were making normal school progress. It appears probable that *either* the reading difficulty is the primary handicap which then leads to secondary antisocial problems, *or* that both the psychiatric and educational difficulties stem from a third group of factors evident in the child and his family from an early stage in the child's development.

---

* A three-year cohort was used for this purpose so that the population studied was 3,300.

REFERENCES.

NEALE, M. D. (1958). *Neale Analysis of Reading Ability Manual.* London : Macmillan.

RUTTER, M. (1965). Classification and categorization in child psychiatry. *J. Child Psychol. Psychiat.*, 6, 71-83.

RUTTER, M. (1967). A children's behaviour questionnaire for completion by teachers ; preliminary findings. *J. Child Psychol. Psychiat.*, 8, 1-11.

RUTTER, M., and GRAHAM, P. (1966). Psychiatric disorder in 10 and 11-year-old children. *Proc. Roy. Soc. Med.*, 59, 382-387.

RUTTER, M., and GRAHAM, P. (1967). The reliability and validity of the psychiatric assessment of the child : (1) Interview with the child ; (2) Interview with the mother. *Brit. J. Psychiat.* (in press).

RUTTER, M., YULE, W., TIZARD, J., and GRAHAM, P. (1966). Severe reading retardation : its relationship to maladjustment, epilepsy and neurological disorders. In *What is Special Education?* Proc. Assoc. for Special Educ., International Conference. London : Association for Special Education.

YULE, W. (1967). Predicting reading ages on Neale's Analysis of Reading Ability. *Brit. J. Educ. Psychol.*, 37, 252-255.

QUESTIONNAIRE MEASURES OF MALADJUSTMENT : A POST-SCRIPT TO THE SYMPOSIUM

BY JACK TIZARD

(University of London Institute of Education)

THE Underwood Committee (1955) and the Scottish Education Department's *Committee on the Ascertainment of Maladjusted Children* (1964) found it impossible to estimate either the prevalence of maladjustment in school children or the unmet needs for services. But without guide-lines as to numbers, any rational planning of services is impossible (Tizard, 1966). It is good, therefore, that psychologists and psychiatrists are again taking up the study of epidemiological problems of mental ill-health in children—problems first explored in this country by Sir Cyril Burt (1933).

Among the questions raised in the discussion of the papers presented to the Cambridge Symposium, three have been singled out for comment :

(1) How should we classify the psychiatric disorders of childhood? Stott's classification appeared to most of the participants in the Symposium to have no advantages over the more widely used Child Guidance classification (Underwood, 1955) or the empirically-based classification outlined by Rutter (1965). The introduction of terms such as 'unsettled,' 'inconsequential,' 'restless' etc., implies that certain reported symptoms consitute a syndrome which is in some way different from or more precisely defined than the types of behaviour described under such terms as 'sthenic neurosis' (Burt, 1933), 'con-

duct disturbance' (Beller, 1962; Cameron, 1955), 'acting out behaviour' (U.S. literature, *passim*), or 'behaviour disorder' (usual Child Guidance classification). This has not been established and the mere proliferation of terms has nothing to recommend it. (Incidentally, a W.H.O. Working Party is at present discussing classification in child psychiatry, and their report, which is likely to be produced in 1968, should do something to bring order and standardisation into the terminology.)

(2) Does the *Bristol Social Adjustment Guide* offer advantages over other inventories? The Guide is widely used, and is said to be liked by teachers. On the other hand, it is expensive to buy and tedious to score in large numbers. Vernon (1964) has given a balanced appraisal of its usefulness and limitations: "To the psychometrist, Stott's instruments are likely to seem distressingly amateurish; but their virtue lies in their simplicity—that they use terms to describe deviant patterns which lay observers can readily apply, and that the psychologist or psychiatrist is presented with an overview, only roughly categorised, so that he can further explore particular symptoms. No doubt the categories overlap a good deal, and might be purified by more thorough statistical analysis. Doubtless also, despite Stott's claim for high inter-rater agreement, there is still a good deal of subjectivity and halo in a typical record as filled in by, say, a teacher."

As there is virtually no information about the comparative advantages of any one of the currently used questionnaires and inventories over any other, psychologists looking for a screening measure might do well to consider whether their needs might not be equally well served by inventories which have not been copyrighted. Among these are Bowlby's (1956), Mulligan's (1964) and Rutter's (1966).

(3) How valid are inventories anyway as screening devices? This is the central question. It has been little studied. However, both Mulligan and Douglas (1963) and Rutter and his colleagues (Rutter and Graham, 1966; Yule and Rutter, *supra*; Graham and Rutter, personal communication) have provided data bearing on this question, and for the record, their findings are summarised below.

Mulligan and his colleagues had teachers complete questionnnaires in respect of 13-year-old children who were attending Child Guidance Clinics, and who had been diagnosed by psychiatrists as suffering from an aggressive and troublesome disorder, or a nervous or neurotic disorder. They compared the findings obtained on these children with those obtained when the same teachers were asked to report on matched controls. The findings are summarised in Table 1.

TABLE 1

MALADJUSTMENT CLASSIFICATION—BASED ON TEACHER RATINGS—OF MALE CHILD GUIDANCE CHILDREN AND THEIR CONTROLS. RESULTS IN PERCENTAGES.

Classification based on teacher ratings	Child Guidance Children			Controls
	Nervous N=50 %	Aggressive N=40 %	Total N=96 %	N=96 %
Nervous	56	6	32	18
Aggressive	4	63	32	9
Mixed	14	9	12	2
Normal	26	22	24	71

About a quarter (24 per cent.) of the maladjusted children were wrongly classified as normal on the basis of the inventory data, while 29 per cent. of normal children were wrongly classified as maladjusted. As maladjusted children constitute only 5–10 per cent. of the general population, the screening power of this inventory still leaves a good deal to be desired. Let us, for example, assume (*a*) that 6·6 per cent. of children are maladjusted ; (*b*) that 75 per cent. of these will be correctly identified on the basis of information obtained from Mulligan's inventory ; (*c*) that 30 per cent. of normal children are misclassified as maladjusted on the basis of inventory scores. The results which would be obtained, if 1,000 children were screened with the help of this inventory, are as in Table 2.

TABLE 2

Inventory Diagnosis		Final Diagnosis	
		Normal	Maladjusted
Normal	670	654	16
Maladjusted	330	280	50
Total	1000	934	66

In correctly identifying fifty maladjusted children (and missing sixteen) one would also classify as maladjusted 280 normal children.

The Isle of Wight findings, if only because they are epidemiologically based and because they do not rely on clinic cases for the validation data, give information which bears more directly on our question. Of 2,193 children screened by both parent questionnaire and teacher questionnaire, both of which had been previously piloted extensively (Rutter, 1967), 157 were selected as ' maladjusted' on the basis of the teacher questionnaire, and 133 were selected as ' maladjusted ' on the basis of the parent questionnaire. Only nineteen children were selected on the basis of both teacher and parent questionnaires. Virtually all of these selected children were seen individually by a child psychiatrist, and much additional information about them and their families was obtained through records and home interviews with parents (Rutter and Graham, 1966 ; Yule and Rutter, *supra*). On the basis of all the available information, a *final diagnosis* was arrived at in respect of each child.

Of the 157 children selected on the basis of the teacher questionnaire, sixty-four were finally diagnosed as maladjusted. The corresponding figures for the other groups were : parent questionnaire, 133 selected, sixty-six finally diagnosed ; both questionnaires, nineteen and fourteen.

Assuming the final diagnosis to be valid, these data give us both the numbers of maladjusted children who were correctly identified as such by each questionnaire (' true positives '), and the numbers of children finally diagnosed as normal who were *incorrectly* judged to be maladjusted on the questionnaires (' false positives ').

An estimate was also made of the number of ' false negatives '—that is, the numbers of maladjusted children in the population who were not identified as such on either questionnaire. A group of 108 physically handicapped children in the same age group served as the validation group for this purpose. These children were selected for special study on physical grounds (through school medical records, hospital records, etc.), and were also studied by questionnaire and by psychiatric examination in virtually the same way as were the other selected children. There were twenty-four who were finally diagnosed as maladjusted, and for twenty of these twenty-four there was questionnaire information available on both teacher and parent inventories. In sixteen cases (80 per cent.) the inventory scores were above the maladjustment cut-off point on one or other of the questionnaires. Thus, the proportion of false negatives was 20 per cent., or to put it another way, for every sixteen children correctly identified by inventory, four children were missed.

In Table 3 the data are expressed as rates per thousand, as was done in Table 2. The prevalence of maladjustment is once again presumed to be 6·6 per cent.

TABLE 3

Inventory Diagnosis		Final Diagnosis	
		Normal	Maladjusted
Normal	876	863	13
Maladjusted :			
(i) on Teacher Inventory	71	42	29
(ii) on Parent Inventory	61	31	30
(iii) on *both* Inventories	8	2	6
(iv) on *either* Inventory	124	71	53
Totals (Normal + those selected on either inventory)	1000	934	66

Assuming the 'final diagnosis' to be valid, the Table shows (*a*) that both the teacher inventory and the parent inventory select about equal numbers of maladjusted children (true positives). The teacher inventory selects forty-two false positives for every twenty-nine true positives, and the parent inventory selects thirty-one false positives for every thirty true positives ; (*b*) that each questionnaire taken by itself selects under half the total number of maladjusted children in the community : (*c*) since, however, each questionnaire picks out different children (the overlap is surprisingly small), when used in combination they pick out four-fifths of the true positives (fifty-three out of an estimated sixty-six) together with one-and-a-third times as many false positives (seventy-one in the present instance compared with fifty-three true positives).

CONCLUSION.

It has yet to be established that any other screening devices will be more efficient than these, and the data from the Isle of Wight study and from Mulligan's both indicate that teacher questionnaires alone cannot be regarded as valid indices of maladjustment. However, teacher questionnaires and parent questionnaires, used in combination to select children for special examination, can cut down very markedly the numbers requiring to be seen. If such questionnaires have been carefully piloted, *and if correct cut-off points are used* (Meehl, 1955) rather few 'cases' need be missed, though large numbers of false positives will be obtained. The evidence thus indicates that inventories can be used very effectively as screening devices but that they are not in themselves sufficient for epidemiological or clinical work.

REFERENCES

BELLER, E. K. (1962). *Clinical Process: A New Approach to the Organisation and Assessment of Clinical Data.* Ill.: Free Press of Glencoe.

BOWLBY, J. (1956). The effects of mother-child separation: a follow-up study. *Brit. J. Med. Psychol.*, 29, 211.

BURT, C. E. (1933). *The Subnormal Mind.* London: Oxford Univ. Press.

CAMERON, K. (1955). Diagnostic categories in child psychiatry. *Brit. J. Med. Psychol.*, 28, 67-71.

MEEHL, P. E. and ROSEN, A. (1955). Antecedent probability of the efficiency of psychometric signs, patterns or cutting scores. *Psych. Bull.*, 52, 194–216.

MULLIGAN, G. (1964). Unpublished Ph.D. thesis, University of London.

MULLIGAN, G., DOUGLAS, J. W. B., HAMMOND, W. A., and TIZARD, J. (1963): Delinquency and symptoms of maladjustment: the findings of a longitudinal study. *Proc. R.S.M.*, 56, 12, 1083-1086.

RUTTER, M. (1965). Classification and categorisation in child psychiatry. *J. Child Psychol. Psychiat.*, 6, 71-83.

RUTTER, M. (1966). A children's behaviour questionnaire for completion by teacher: preliminary findings. *J. Child Psychol. Psychiat.*, 8, 1-11.

RUTTER, M., and GRAHAM, P. (1966). Psychiatric disorder in 10 and 11-year-old children *Proc. R.S.M.*, 59, 382-387.

SCOTTISH EDUCATION DEPARTMENT (1964). *Ascertainment of Maladjusted Children: Report of the Working Party appointed by the Secretary of State for Scotland.* Edinburgh: H.M.S.O.

TIZARD, J. (1966). Mental subnormality and child psychiatry. *J. Child Psychol. Psychiat.*, 7, 1-15.

UNDERWOOD COMMITTEE (1955). *Report of the Committee on Maladjusted Children.* London: H.M.S.O.

VERNON, P. E. (1964). *Personality Assessment: A Critical Survey.* London: Methuen.

(*Manuscripts submitted 30th September and 13th November*, 1967)

RATER-RATEE PERSONALITY IN THE ASSESSMENT OF TEACHING ABILITY

By K. B. START

(School of Education, University of Manchester)

SUMMARY. The members of staff of an English secondary modern school completed a battery of tests, one of which was the 16 Personality Factor Test of Cattell. The head teacher completed the test battery as well as rating each assistant teacher on nine aspects of teaching competence. These ratings were highly inter-correlated and on factor analysis, two disproportionate principal components emerged. Scores for the similarity of the personality profiles of head teacher and teacher were calculated. These were then compared with the head teacher's rating of the teachers' teaching ability, of their relations with the head teacher and of their relations with staff colleagues. It appeared that teachers with profiles very similar to, or very different from that of the head teacher, usually had higher ratings for teaching ability and for social competence. The teachers with the lowest ratings for teaching had a tougher, self-confident unconventionality which would tend to make them independent of the head teacher.

I.—INTRODUCTION.

HEAD teacher's assessments of assistant teachers form a major part of the latter's applications for other teaching posts, whether these are promotions or not. Such assessments by head teachers have obvious professional significance. They take on additional importance when they are used as criteria in research studies attempting to evaluate training courses or estimating the psychological and social factors which are preferentially found in 'good' teachers. The criterion problem reaches a more involved level when teachers are matched to evaluate particular teaching methods.

The face validity of the assessment would appear high—the judgment of the professional competency of one person by another who is assumed to be experienced and proficient in the skills to be assessed and whose personal advancement has been through the existing professional system. However, one person's judgment of another is notoriously suspect, either individually (Sarbin, Taft and Bailey, 1960) or collectively as in an interview by committee (Eysenck, 1955).

A recent study (Wiseman and Start, 1965) has shown that teaching ability as described in the head teacher's reference was not significantly correlated with teaching practice grades given five years earlier by training college staffs. From this it could be inferred that the teachers changed considerably between ratings, or that the reliability of the rating procedure was low, or that the college staff and the head teachers used different criteria. More needs to be known about the bases on which these ratings were made.

II.—METHOD.

An analysis of 248 head teachers' references had suggested that there were a number of broad areas on which the head teachers tended to comment. In the present study the head teacher was asked to use a five-point scale to rate each staff member on nine different criteria, teaching ability, relations with colleagues, class control, relations with parents,administrative ability, relations with children, intelligence, extra-mural contributions and relations with the head teacher. The head teacher was asked to select the five out-standing members of staff on each trait (scores 5) and the five weakest (scores 1). The remaining teachers were assessed average, above average and below average (scored 3, 4 and 2, respectively). The head teacher was asked to keep the number of teachers in each of the five groups within certain limits so that roughly a normal distribution was obtained. The rating instruction booklet also contained the warning " *Rate every member of staff on one trait before moving to a second trait for any one teacher*." Nine rating sheets were provided, one for each of the traits listed.

The staff (N=39) of a secondary modern school were asked to co-operate by completing a test battery which included the 16 P.F. Test (Form B). The head teacher allocated a number to each member of staff who returned the test material sealed in a numbered envelope to the investigator. The head teacher also completed a set of the test material and provided the teachers' ratings which were compared with test scores by means of those numbers.

III.—RESULTS.

Four members of staff did not co-operate but as they had been rated 5, 4, 3 and 1 for *teaching ability* and 5, 4, 2 and 1 for relations with head teacher, there seems little evidence for a systematic bias on these ratings. However, they were rated 3, 2, 2 and 1, in relations with colleagues which might suggest that they were less friendly with their colleagues than were other members of staff.

From the analysis of the head teacher's ratings of all thirty-eight teachers, it was obvious that the ratings of each teacher on the nine variables were highly correlated. Whilst the lowest inter-correlation was ·249 (staff relations with administrative ability), twenty-nine of the thirty-six inter-correlations exceeded ·5. Obviously, a considerable halo effect was occurring. The sum of the ratings from the previous nine ratings was used as a tenth variable and the data subjected to factor analysis. Two principal components were revealed, a major one, describing 60·2 per cent. of the variance and a minor one, 9·5 per cent. Details of the loadings appear in the first part of Table 1.

TABLE 1

VARIABLE LOADINGS ON FACTOR ANALYSIS OF HEAD TEACHER'S RATINGS.

Variables	Principal Components		Rotated Factors		
	PC1	PC22	RF1	RF2	HH
Teaching ability	·922	·108	·576	·728	·86
Relations with staff	·711	·618	·066	·939	·89
Class control	·851	—·291	·808	·396	·81
Relations with parents	·882	—·273	·817	·430	·85
Administrative ability	·687	—·552	·876	·095	·78
Relations with children	·858	·012	·559	·615	·74
Intelligence	·693	·228	·329	·651	·53
Extra-Mural contribution	·760	—·013	·547	·529	·58
Relations with Head	·891	·190	·496	·765	·83
Total	1·000	·000	·707	·707	1·00
% Variance	69·18	9·45	39·3	39·3	—

The comprehensive nature of the first component (PC1) is obvious. The polarity of PC2 is interesting, in that it distinguishes *administrative ability* from *relations with staff*. It would seem that being thought a good administrator might reduce the cordiality of one's relations with one's colleagues. Alternatively, it might reflect the status difference between these with administrative ability (already in positions of responsibility ?) and the younger members of the staff room.

With an eigen value of ·945, the minor factor was slightly below the value usually taken as significant but it was included for rotation calculations (Varimax and graphically). The variance was recast around the foci of *administrative ability* (RF1) and *relations with staff* (RF2), each with 39·3 per cent. of the total variance.

It is interesting to observe on which of the rotated factors certain of the variables preferentially. The variances of *total, extra-mural activity* and *relations with children* are shared equally. The larger part of the variance of *relations with parents* and *class control* occur with *administrative ability* in RF1. On the other hand the variables of *teaching ability, relations with Head* and *intelligence* associate the higher share of their variance with *relations with staff* in RF2. Unfortunately, the low sample numbers reduce the reliability of this picture.

With such an obvious halo the composite rating might seem the appropriate one to use as the criterion. Professionally, however, it is the estimate of teaching ability which has face validity. Personal relations may cloud professional judgment and thus the rating for *relationship with the head teacher* was of interest. The rating of the cordiality of *relationships with colleagues* was a second estimate of the teacher's social competence.

Similarity of individual personality factors.

A comparison of the head teacher's 16 P.F. profile, with the mean profiles for teachers rated highest (5) and lowest (1) on teaching abilities indicated that the major differences in the profiles fell into two categories : first, those in factors A, B, L and M (in the 16 P.F. profiles) where the mean profile of the teachers rated 1 (low) for teaching ability differed considerably from the other two profiles ; and second, in factors F, H, Q1 and Q2 where the difference lay between the profile of the Head and the two extreme groups. Thus, the only occasions where scores of the head teacher and teachers rated lowest on teaching ability appeared to differ from teachers rated highest on teaching ability were on factors Q4 and G. In endeavouring to establish the significance of these differences the problem of low numbers and relatively high standard deviation arose. As a result, when *sex* was partialled out only L– and M were significantly related to teaching ability although B and H were close to significance.

Using the head teacher's personal score on a particular personality factor as focus, the eleven assistant teachers with the closest scores to the Head's on that factor were noted. This group was taken to be *similar* to the Head on this factor and the mean scores of this group for each of the three rated competences were calculated. Teachers with factor scores above that of this group were considered as the *high* group and similarly teachers scoring below the range of the *similar* group were taken as members of the *low* group. For each of these two extreme groups, mean ratings of teaching ability, of relations with Head and of relations with colleagues were calculated. This procedure was repeated for each of the sixteen factors.

When the Head's personal score was at or about the median of the total group, three approximately equal sub-groups were available as for Factor N. However, this balance proved the exception rather than the rule. When the Head's score was near either limit of the total group scores, then only two groups resulted, the Head's and a higher, as in Q4 for example, or the Head's and a lower group as with Q2. Within the sixteen factors examined, numbers in sub-group varied considerably.

On inspection, in five of these traits, groups of teachers with higher or lower scores than the head teacher's personal score had a lower mean competence-rating from the head teacher. Similarity and high rating were associated for general intelligence (B), emotional stability (C), dominance (E), unconventionality (M) and with a similar level of self-sentiment (Q3). Few of these trends, however, were significant. In Table 2 are the factors on which significant differences between sub-group mean ratings of competence were found by critical ratio methods.

TABLE 2

16 P.F. FACTORS ON WHICH HEAD TEACHER-TEACHER SIMILARITY WAS ASSOCIATED WITH SIGNIFICANTLY DIFFERENT RATINGS OF TEACHER'S COMPETENCE.

16 P.F. Factor	Teacher's score in relation to Head's score	N	Mean Rating		
			Teaching Ability	Relations with Head	Relations with Colleagues
B General Intelligence	High	—	—	—	—
	Similar	11	3·91	4·00	4·09
	Low	23	2·70**	3·09*	3·04**
G Conscientious Persistent	High	3	3·00	4·00	4·33
	Similar	12	2·58	3·08	3·17
	Low	19	3·42*	3·47	3·67
L Suspecting Jealous	High	20	2·60**	2·90	2·85
	Similar	11	3·73] *	3·91*] *	4·091**] **
	Low	3	4·00	4·67	4·33
Q1 Radical	High	4	2·50	3·25	3·50
	Similar	11	2·82] *	3·00	3·09
	Low	19	3·37	3·63*	
Q2 Self-sufficient	High	—	—	—	—
	Similar	11	2·46	2·73	3·00
	Low	23	3·39	3·70*	3·57

Difference between means significant at 5 per cent. (*) and 1 per cent. (**) levels.

Overall it would appear that the teacher rated successful in this school was rather like the head teacher but somewhat more academic (A – and B +), less sure of himself (G –, Q1 and Q2), more accommodating (L –) and not as shrewd (N—). The head teacher saw the successful teacher, therefore, as a more intellectual, less confident or less experienced version of himself, and one who has to draw from the head teacher's experience in the field. On the other hand, the teachers rated poorer appear to be more confident, tougher, shrewder, tenser, more radical and more unconventional than the head teacher. The picture of the good teacher clarifies as a somewhat more intellectual self-image with acceptable dependency on the present head teacher's experience. The teacher seen as less capable had a tough, radical, self-confident, unconventionality which would tend to make him independent of the Head and rarely seek his counsel. However, with such a small group, these tentative interpretations can best be regarded as hypotheses to be examined by a more detailed analysis of a much larger group, perhaps using interview techniques. Other variables might impinge on this situation, perhaps personal appearance, teaching experience, age, area of teaching subject and 'social class.'

The social relations ratings appear to follow closely the teaching efficiency estimate, as might be expected from the halo effect indicated by the Factor Analysis. Actually, the specific inter-correlations were: *teaching ability* with *relations with colleagues* ·78, and with *relations with head teacher* ·82. The correlation between the two ratings of social relations was ·78.

Similarity of personality profiles (by grouping).

A teacher could, by the method described earlier, fall into the *similar* group for each of the sixteen factors. Teachers, therefore, could be given a *similarity* score based on the number of times they fell into the *similar* group. The range was found to be 2–10. The ratings of *Teaching Ability* of the members of each *similarity* sub-group were then noted.

Although there were consistent trends for the six teachers with personality profiles closest to that of the head teacher to be rated high on aspects of their professional and social competence, not one of the differences between means of the sub-groups reached the minimum level usually taken for statistical significance.

Similarity of personality profiles (by mean squares).

The second estimate of similarity (actually dissimilarity) used was that outlined in the 16 P.F. Handbook and basically is the sum of the squares of the difference between the scores on each personality factor of the Head and the assistant teachers. This approach made it possible to observe how factors differed in the approximation of individual scores to that of the head teacher Greatest convergence on the head teacher's score were scores on Factor I and least on Factor H, which had sums of squares of differences of 58 and 525 respectively. The remaining fourteen traits had differences between 100 and 200.

An overall dissimilarity score for an assistant teacher's 16 P.F. profile with that of the head teacher was calculated by summing the squared sten difference on each trait and this estimate was then compared with ratings of *teaching ability*, *relations with head teacher* and *relations with staff*. Dissimilarity scores ranged from 25 to 147 (mean 93·5, s.d. 30·1) and correlated negatively, but not significantly with teaching ability, with relations with staff and with relations with head teacher.

The dissimilarity scores were ordered and five groups based on a normal distribution were considered (N: 4, 8, 10, 8, 4). The differences in the mean ratings of these groups for *teaching ability*, *relations with colleagues* and *relations with head teacher* were then tested by means of the critical ratio technique.

It appeared that the most similar group had a significantly (5 per cent.) higher mean rating for *teaching ability* than all but the least similar group. The latter had a mean rating higher than those of the three middle groups which failed to be significantly better (5 per cent.) than the median group only.

As regards the *relations with colleagues* a similar pattern was found except that significant differences occurred only between the most similar and the group on either side of the median group. Mean scores for relations with Head were in the same order, but in this case not one of the differences was significant.

It was notable that the highest means were those of the two extreme groups. This curvilinear relation had also been observed in the detailed inspection of individual factors, particularly, A, G and O. To see if there was any common ground between those two extreme groups, the individual 16 P.F. profiles of the members of the groups were compared.

Among the first order personality factors an obvious similarity was noted in *I* with all scores being 8 or 9 (a range of 2 stens). However, fifteen other assistant teachers had *I* scores in this range, and these had mean scores of 2·53 for teaching ability, 3·07 for relations with staff and 2·93 for relations with Head. The scores of these eight extreme teachers had a range of 4 stens on Factors B and O and 5 stens on Factors F and Q4, but the number of teachers with scores within this range on these factors was 22, 19, 21 and 19, respectively.

A final comparison was made graphically by comparing the four 16 P.F. profiles of the most and least similar assistant teachers rated 5 (high) for teaching ability and the most and least similar teachers rated 1 (low) for teaching ability. The two high rated teachers had identical scores on E, F, I and L. They also appear to differ from the low scoring teachers on these factors and perhaps G.

IV.—DISCUSSION.

The curvilinear relation between the similarity of the teacher's and the head teacher's personality and the head teacher's rating of the assistant teacher's professional competence is intriguing. Perhaps it is an over-compensation by the Head in rating people with whom he has little in common. Perhaps with such completely different personalities, these teachers achieve success in situations where the head teacher has failed and are respected on this account. Both extreme groups might have had high score on *I* in common, but this was difficult to establish within this sample. It would seem, therefore, that this head teacher would pick as good teachers people with very similar or very dissimilar personalities to his own. It may not be inherently wrong for the success of teachers to be related to the similarity of their personality with that of the head teacher, provided two things can be shown: firstly, that the head teacher is a validated measuring instrument, i.e., a successful teacher himself; and secondly, that similarities essential to good teaching can be assessed by personality variables.

In the absence of an objective measure of 'success' in teaching one is forced back on pragmatic criterion such as promotion in which the assumption is that competency has been rewarded. The weaknesses of this criterion are patent, but in the absence of professionally acceptable measures, such as pupil-gain, there is not a great deal of choice. In these circumstances, whatever other attributes aided their promotion, head teachers must be assumed to be at least competent teachers.

There have been many studies which have shown that personality traits can have low but significant correlations with assessments of teaching competence. Getzels and Jackson said "the studies . . . are too numerous for individual mention" (1963, p. 506), but go on to mention all the American work in this field in their chapter on "The Teacher's Personality and Characteristics." The question whether head teachers are competent classroom teachers remains unanswered. Head teachers' ratings of the professional competence of assistant teachers may be associated with the similarity in personalities of the rater and ratee, but even so, such assessments can be crudely rationalised in the absence of unimpeachable objective measures of this professional ability. It may not be a good estimate, but it may be acceptable in the present circumstances. However, the use of a measure such as this obviously leads to certain predictable problems. If the head teacher's personality is a significant feature of a teacher's rating, then for no other reason than that of different head teachers, the competence of an assistant teacher is likely to appear to vary from school to school. This, coupled with the varying teaching experiences that teachers gather in the passage of time, and the teacher's ability to benefit from such experiences, may cause ratings of competence to fluctuate widely. Perhaps this is part of the explanation of the findings of studies which have found little or no correlation between assessments during college courses and subsequent head teacher's ratings. If, as Langevelds (1963) suggested, "the teacher's type is thus a primary character type which has received accentuation as the result of training and occupation" (p. 405) and if the estimate of the efficiency of the professional type is determined by head teachers' assessment, it may be that "general characterology . . . in which professional type is not assimilated" (p. 405) may be more important, not less important as he suggests.

Acknowledgements.—The considerable contribution of Miss Patricia Clayton in the scoring and collation of the data in this study is gratefully acknowledged.

V.—References.

Cattell, R. N., and Eber, H. W. (1957). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign (Illinois): I.P.A.T.

Eysenck, H. J. (1955). *Uses and Abuses of Psychology*. London: Pelican.

Getzels, J. W., and Jackson, P. W. (1963) in N. L. Gage (Editor) *Handbook of Research on Teaching*. Chicago: Rand McNally.

Langeveld, M. J. (1963) in Bereday, G. Z. F. and Lauwerys, J. A. (Editors). *The Education and Training of Teachers*. (Yearbook of Education, 1963). London: Evans Bros.).

Sarbin, T. R., Taft, R., and Bailey, D. E. (1960). *Clinical Inference and Cognitive Theory*. New York: Holt, Rinehart and Winston, pp. 212-225.

Start, K. B. (1966). The relation of teaching ability to measures of personality. *Brit. J. Educ. Psychol.*, 36, 159-165.

Wiseman, S., and Start, K. B. (1965). A follow-up of teachers five years after completing their training. *Brit. J. Educ. Psychol.*, 35, 342-361.

(*Manuscript received 14th November*, 1966)

TEST CONSTRUCTION—A TECHNIQUE FOR PRODUCING AND VERIFYING A NORMAL DISTRIBUTION OF ITEMS

BY NORMAN FRANCE

Faculty of Education, University of Saskatchewan, Regina Campus)

SUMMARY. Suggestions are offered for facilitating the standardization and interpretation of test scores. By using a table for transformating percentiles into equivalent normalized standard scores it is possible to apply a simple approximate test for a normal distribution. If the graph showing the relation of the raw scores to the normalized standard score is a straight line, then the raw scores are themselves normally distributed. The same table enables the order of difficulty of test items to be more precisely described so that a selection of items can be made which will produce a normal distribution of scores in the final form of the test.

I.—INTRODUCTION.

THE interpretation of test scores is easier if the raw scores and the corresponding standardized scores are normally distributed. The procedures described in this note have proved to be simple and effective in actual use. It will be seen that neither computer nor electric calculator is really necessary, although, if large numbers of testees are involved, a punched card sorting procedure can be useful.

The use of a ' normalized standard scale ' derived from percentiles, instead of 'deviation scores,' expressed as the ratio of the raw deviations to the standard deviation, originated with Karl Pearson (1902). Table II in his *Tables for Statisticians and Biometricians* (1914) was prepared for this purpose. The technique was first used in educational and psychological testing by Thorndike (*Mental and Social Measurements*, 1904). His Tables XXXI—XXXII, derived from Pearson's, allow for the effects of the differing size of the sample.

II.—THE TEST FOR A NORMAL DISTRIBUTION.

If a set of scores is normally distributed, the percentiles are related to the standard scores as shown in Table 1. ' Standard score ' is defined here as

$$\frac{\text{Raw score—Mean}}{\text{Standard Deviation}}.$$

TABLE 1

RELATION BETWEEN PERCENTILES AND STANDARD SCORES.*

Percentile	Standard Score	Percentile	Standard Score
(99)	+2·33	(1)	—2·33
90	+1·28	10	—1·28
80	+0·84	20	—0·84
70	+0·52	30	—0·52
60	+0·25	40	—0·25
50	0·00	50	0·00

* *Source* : K. Pearson (1914).

[illegible] Psy[illegible] Research
[illegible]
10.3.81

The most common standardization for tests in Great Britain has a mean of 100 and a standard deviation of 15. This is much easier to use in practice and is more readily understood in the classroom than standard scores with a mean of 0 and a standard deviation of 1. The idea of avoiding negative marks by choosing a mean such as 50 or 100 and the use of a standard deviation of 10 or 15 was developed around 1915. In the United States, the best known example is McCall's ' T-Scale.' (1922).

The choice of a mean of 100 and a standard deviation of 15 in Britain follows from the early observation that the follow-up of London children tested with the Binet scale revealed a greater constancy for the percentile rank than for the intelligence quotient calculated as a ratio of mental to chronological age. As a result, tables were published for converting ' mental ages ' and ' mental ratios ' to percentiles and *vice versa*. Previous workers, such as Galton and Thorndike had used the probable error or ' semi-interquartile range ' as the unit. With the English translation of the Binet scale this was usually 10 I.Q. points, i.e., half the age group fell between 90 and 110 I.Q. On changing to the standard deviation as unit, this became equivalent to 15 I.Q. points. This procedure was consequently applied to attainment as well as to intelligence testing. More recently, in America, it has become customary to take 16 I.Q. points as the standard deviation and this is referred to as the ' deviation I.Q. ' (Cronbach, *Essentials of Psychological Testing* (1960, page 171). Wechsler, however, prefers 15 I.Q. points. Perhaps the best known table " For Transmuting Per Cent. Positions to a Linear Scale " is Hull's (*J. App. Psychol.*, VI, p. 385) and this has been adapted by many writers as, for example, J. P .Guilford (1965, page 520). Table 1 can now be expanded into a ready reckoner (Table 2) which will provide the standardized score equivalents of each percentile level for a normal distribution with a mean of 100 and a standard deviation of 15.

A frequency distribution of scores to be tested for normality is first converted to a cumulative frequency distribution, then to percentiles, and by using Table 2, these can be transformed to the corresponding normalized standard scores. If, when the raw scores are plotted against normalized standard scores, the resulting graph is a straight line, then the original scores are distributed normally.

The use of a graphic method as a rough test of normality is, of course, not new. The use of arithmetical probability graph paper enables a graph to be drawn for the relationship between scores and corresponding cumulative percentiles without special calculation. With this graph paper, a normal distribution produces a straight line. Table 2 produces the same result without the need for special graph paper. For many practical purposes, it may be sufficient to check the agreement with normality by this graphical method. It is, of course, desirable to apply a chi-squared test if more than a coarse approximation is required.

For a normal distribution it should be noted that the gradient of the straight line on the graph is proportional to the standard deviation. If two distributions are represented by parallel lines, they have the same standard deviation. If two distributions intersect at a point for which the abscissa is 100, they have the same mean.

Table 2 should be readily understood by anyone accustomed to using four-figure logarithm tables. It emphasizes the symmetrical nature of the distribution without requiring the user to subtract from 100 mentally. If the size of the sample is less than 30 or thereabouts, some care will be needed in interpretation. For larger samples up to about 300 or so, the use of percentiles to the nearest whole number will usually be sufficient and the columns for decimal percentile additions can be ignored. The full table is perhaps of the greatest value for samples of 1,000 or more.

TABLE 2

CONVERSION OF PERCENTILES TO NORMALIZED STANDARD SCORES.
(M=100 S.D.=15)

Percentile	Standard Score	Percentile	Standard Score	Additions to Standard Score for Decimal Percentile Additions								
				·1	·2	·3	·4	·5	·6	·7	·8	·9
*0	40·0	—	—	14·4	16·8	18·8	20·2	21·4	22·3	23·1	23·9	24·5
	—	99	134·9	0·6	1·2	2·0	2·8	3·7	4·9	6·3	8·3	10·7
1	65·1	—	—	0·5	1·0	1·5	1·9	2·4	2·7	3·1	3·4	3·9
	—	98	130·8	·3	·7	1·0	1·4	1·8	2·2	2·6	3·1	3·6
2	69·2	97	128·2	·3	·5	·8	1·1	1·3	1·6	1·8	2·0	2·3
3	71·8	96	126·3	·2	·4	·6	·7	1·1	·9	1·3	1·5	1·7
4	73·7	95	124·7	·2	·3	·5	·6	·8	·9	1·1	1·3	1·4
5	75·3	94	123·3	·2	·3	·4	·6	·7	·8	1·0	1·1	1·2
6	76·7	93	122·1	·2	·3	·4	·5	·6	·8	·9	1·0	1·1
7	77·9	92	121·1	·1	·2	·3	·4	·5	·6	·7	·9	1·0
8	78·9	91	120·1	·1	·2	·3	·4	·5	·6	·7	·8	·9
9	79·9	90	119·2	·1	·2	·3	·4	·5	·5	·6	·7	·8
10	80·8	89	118·4	·1	·2	·3	·4	·4	·5	·6	·6	·7
11	81·6	88	171·6	·1	·2	·3	·3	·4	·4	·5	·6	·7
12	82·4	87	116·9	·1	·2	·2	·3	·3	·4	·5	·5	·6
13	83·1	86	116·2	·1	·2	·2	·3	·3	·4	·5	·5	·6
14	83·8	85	115·5	·1	·2	·2	·3	·3	·4	·5	·5	·6
15	84·5	84	114·9	·1	·2	·2	·3	·3	·4	·5	·5	·6
16	85·1	83	114·3	·1	·1	·2	·3	·3	·4	·5	·5	·5
17	85·7	82	113·7	·1	·1	·2	·3	·3	·4	·4	·5	·5
18	86·3	81	113·2	·1	·1	·2	·3	·3	·4	·4	·5	·5
19	86·8	80	112·6	·1	·1	·2	·3	·3	·4	·4	·5	·5
20	87·4	79	112·1	·1	·1	·2	·3	·3	·4	·4	·5	·5
21	87·9	78	111·6	·1	·1	·2	·3	·3	·4	·4	·5	·5
22	88·4	77	111·1	·1	·1	·2	·2	·2	·3	·3	·4	·4
23	88·9	76	110·6	·1	·1	·2	·2	·2	·3	·3	·4	·4
24	89·4	75	110·1	·1	·1	·2	·2	·2	·3	·3	·4	·4
25	89·9	74	109·7	·1	·1	·1	·2	·2	·3	·3	·4	·4
26	90·3	73	109·2	·0	·1	·1	·2	·2	·3	·3	·4	·4
27	90·8	72	108·8	·0	·1	·1	·2	·2	·3	·3	·3	·4
28	91·2	71	108·3	·0	·1	·1	·2	·2	·2	·3	·3	·3
29	91·7	70	107·9	·0	·1	·1	·2	·2	·2	·3	·3	·3
30	92·1	69	107·4	·0	·1	·1	·2	·2	·2	·3	·3	·3
31	92·6	68	107·0	·0	·1	·1	·2	·2	·2	·3	·3	·3
32	93·0	67	106·6	·0	·1	·1	·2	·2	·2	·3	·3	·3
33	93·4	66	106·2	·0	·1	·1	·2	·2	·2	·3	·3	·3
34	93·8	65	105·8	·0	·1	·1	·2	·2	·2	·3	·3	·3
35	94·2	64	105·4	·0	·1	·1	·2	·2	·2	·3	·3	·3
36	94·6	63	105·0	·0	·1	·1	·2	·2	·2	·3	·3	·3
37	95·0	62	104·6	·0	·1	·1	·2	·2	·2	·3	·3	·3
38	95·4	61	104·2	·0	·1	·1	·2	·2	·2	·3	·3	·3
39	95·8	60	103·8	·0	·1	·1	·2	·2	·2	·3	·3	·3
40	96·2	59	103·4	·0	·1	·1	·2	·2	·2	·3	·3	·3
41	96·6	58	103·0	·0	·1	·1	·1	·2	·2	·3	·3	·3
42	97·0	57	102·6	·0	·1	·1	·1	·2	·2	·3	·3	·3
43	97·4	56	102·3	·0	·1	·1	·1	·2	·2	·3	·3	·3
44	97·7	55	101·9	·0	·1	·1	·1	·2	·2	·3	·3	·3
45	98·1	54	101·5	·0	·1	·1	·1	·2	·2	·3	·3	·3
46	98·5	53	101·1	·0	·1	·1	·1	·2	·2	·3	·3	·3
47	98·9	52	100·8	·0	·1	·1	·1	·2	·2	·3	·3	·3
48	99·2	51	100·4	·0	·1	·1	·1	·2	·2	·3	·3	·3
49	99·6	50	100·0	·0	·1	·1	·1	·2	·2	·3	·3	·3

* Actually 0·003 per cent.

E.g., Percentile Level =30·8 per cent.
Standard Score =92·1+·3
=92·4

Example: The second line of Table 3 shows the distribution of raw scores in a test of Mathematics obtained by 6,872 boys and girls in Grades 6 and 7 (median ages 11+ and 12+, respectively, in a selection of elementary schools in the Province of Quebec in May, 1966). Each line in the table represents a stage in the conversion process.

TABLE 3

CONVERSION OF RAW SCORES TO NORMALISED STANDARD SCORES.

Range of Raw Scores : . .	0–4	5–9	10–14	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	Total
Number of Children	1	16	120	477	1155	1813	1782	1065	375	61	7	6872
Cumulative Frequency . .	1	17	137	614	1769	3582	5364	6429	6804	6865	6872	6872
Percentile	0·0	0·2	2·0	8·9	25·7	52·1	78·1	93·5	99·1	99·9	100·0	100·0
Corresponding Normalized Standard Score (From Table 2)	40·0	56·8	69·2	79·8	90·2	100·8	111·7	122·7	135·5	145·6	160+	—
Increment of Normalized Standard Score	16·8	12·4	10·6	10·4	10·6	10·9	11·0	12·8	10·1	14·4+	—	

The graph when drawn shows the relationship of the normalized standard score to the raw score at the bottom of the next higher interval in the range of raw scores. Thus, for a standard score of 111·7, 78·1 per cent. of the children score less than 35 raw score. The graph is seen to be virtually a straight line, which means that the scores on this Mathematics test are close to normal in their distribution. This can be checked approximately by noting the regularity of the increments in the last row of Table 3, but more precisely by using a chi-squared test. Here chi-square is 10·9 ; with 10 degrees of freedom P is about 0·36. Thus, the actual frequencies are normally distributed to a very close approximation. Comparison with a theoretical normal distribution fitted to the observed percentages shows the actual distribution to have a slight but insignificant flattening in the centre. Between the second and the 94th percentiles, the agreement with a normal distribution is very close. One of the aims here had been to produce a test which would have a normal distribution of scores when administered to English speaking children in grades 6 and 7 in the Province of

Quebec. The evidence suggests that this aim was achieved. The method of selecting items in order to produce such a normal distribution of scores is described in the following section.

III.—CHOICE OF ITEMS TO PRODUCE A NORMAL DISTRIBUTION.

The results of try-out tests were first analysed so as to eliminate items with a large sex-difference, ambiguity or poor discrimination. The percentage of children in Grades 6 and 7 failing to answer each remaining item was used as a basis to allot a normalized standard score by means of Table 2. The items were then arranged in this order of difficulty. The final selection of fifty-six items was made by choosing seven items with a standard score in the range 60 to 69, seven items in the range 70–79, and so on for each range up to 130 and above. The use of normalized standard scores instead of percentages to measure difficulty level has the advantage of achieving a more precise and even distribution in the grading of items in order of difficulty.

IV.—PREPARATION OF THE TABLE FOR CONVERTING RAW SCORES TO NORMALIZED STANDARD SCORES.

The graph relating raw scores to normalized standard scores makes it possible to prepare a conversion table which will give the equivalent normalized standard score (in whole numbers) for each raw score. For this same population of 6,872 Quebec children the converted scores will now form an almost perfect normal distribution. (To save space only even scores are given in Table 4).

TABLE 4

MATHEMATICS TEST—TABLE FOR CONVERTING RAW SCORES TO NORMALIZED STANDARD SCORES.

Raw Score	Normalized Standard Score	Raw Score	Normalized Standard Score
10	57	30	101
12	61	32	105
14	66	34	110
16	70	36	114
18	75	38	119
20	79	40	123
22	83	42	127
24	88	44	132
26	92	46	136
28	97	48	141
30	101	50	145

In Table 4 values for the normalized standard score found from the graph are first inserted for raw scores of 10, 20, 30, 40 and 50. (The value for a raw score of 20 is actually nearer 80 than 79, but this slight approximation ensures even increments of 22 for each 10 points of raw score throughout the whole table.) Interpolation is achieved as is shown in Table 5.

TABLE 5

INTERPOLATION TO PRODUCE INDIVIDUAL ITEMS OF THE CONVERSION TABLE.

Raw Score (10)	11	12	13	14	15	16	17	18	19	(20)
Add to N.St.Sc.	2·2	4·4	6·6	8·8	11·0	13·2	15·4	17·6	19·8	(22·0)
Whole number	2	4	7	9	11	13	15	18	20	(22)

Any normalized standardized score is interpreted as follows: Suppose a child obtains a normalized standard score of 123. Now refer to Table 2. The corresponding percentile level is 93·7. This child's mathematical ability places him just above 93·7 per cent. of a large random sample. Alternatively, he can be regarded as falling just below the most able 6·3 per cent.

V.—CONCLUSION.

(1) A detailed table has been constructed enabling normalized standard scores to be determined directly from the percentile rank. It is, however, accurate only for large samples. This table will be found especially useful:

(*a*) when, as so often happens, the distribution of the test-scores is markedly skewed or departs in other ways from the normal. (In this case a mechanical aid would be helpful when dealing with large samples. A computer programme for mechanically converting (i) scores into percentiles and (ii) percentiles into normalized standard scores would be time saving); and

(*b*) when the data are not expressed in terms of scores, but in ranks or classes which enable percentiles to be determined;

If it is known that the distribution is very near to the normal, then it will usually be easier, quicker, and more accurate to use an equation which calculates standard scores direct from the mean and standard deviation of the raw scores.

(2) When it is necessary to test whether the distribution is approximately normal, the standardized scores determined from the table can be plotted as a graph against the raw scores ;and the line fitting the points should be straight, if the distribution is approximately normal; if curved, it may enable the investigator to judge the most appropriate transformation (logarithmic, square root, or the like). Where the distribution appears to be nearly normal, goodness of fit should be tested by the usual chi-squared procedure.

ACKNOWLEDGEMENT.—The author is much indebted to Sir Cyril Burt for many suggestions which have been incorporated in this paper and in particular for the outline of development of a normalized standard scale in Section II.

VI.—REFERENCES.

PEARSON, KARL (1914). *Tables for Biometricians and Statisticians.*

THORNDIKE, E. L. (1904). *Mental and Social Measurements.* New York: Teachers College.

MCCALL, W. A. (1922). *How to Measure in Education.* New York: Macmillan.

WALKER, H. M., and LEV, J. (1953). *Statistical Inference.* New York: Holt, Rinehart and Winston, Inc.

ANASTASI, A. (1954). *Psychological Testing.* New York: Macmillan.

CRONBACH, L. J. (1960). *Essentials of Psychological Testing.* New York: Harper and Row.

(*Manuscript received* 10*th February*, 1967)

A FACTORIAL STUDY OF COMMUNICATION AND ABILITY IN DEAF SCHOOL LEAVERS

By G. W. G. MONTGOMERY

(*Research Unit, Donaldson's School for the Deaf, Edinburgh*)

SUMMARY. Seventy profoundly deaf school-leavers were rated by their teachers for speech, lip-reading, written work and manual communication, and all but four of them were given a comprehensive battery of nineteen standardised tests of cognitive abilities. Pure tone audiometric responses were also recorded. Intercorrelations of all variables were calculated and the correlation matrix was subjected to a principal components analysis. The resultant general and bipolar factors seemed an inappropriate description of the cognitive abilities tested and graphical rotation was undertaken to obtain more meaningful configurations. Orthogonal factors of non-verbal reasoning, written vocabulary, numerical ability, manual communication and residual hearing were identified. The association of speech development with residual hearing and the absence of significant negative correlations between all variables was noted. Some educational implications of these results are suggested, namely: (1) the techniques for the development of speech in deaf children are notably unsuccessful with the majority of those who have no usable residual hearing; (2) the alleged incompatibility of oral and manual skills is not observable statistically as a negative correlation and (3) the alleged negative correlation between performance test results and verbal attainment does not obtain but the low positive relationship found implies that current educational methods do not effectively gear the intelligence of deaf children to scholastic achievement.

I.—INTRODUCTION.

MUCH of the research into the cognitive abilities of 'the deaf' is vitiated by a loose definition of deafness. In fact, within the larger deaf group defined simply by some impairment of hearing there are many sub-groups with widely different abilities and handicaps (Taylor, 1967). There are, for example, many such differences between the test performance of the congenitally deaf and the person with later 'adventitious' deafness; between those whose hearing impairment is central and those with peripheral impairment; between the partially deaf who may perceive amplified speech and the profoundly deaf person who cannot hear his own voice even with benefit of amplification. These differences in performance are so pronounced, especially in tests dependent upon linguistic development, that any presentation of the results of cognitive tasks, skills or standardised tests which does not allow for, or at least record, the age of onset, type and degree of deafness scarcely deserves to be taken seriously.

This study is concerned only with the prelingually, profoundly deaf, that is those whose deafness occurred before the development of infant speech and whose inability to hear their own voices precludes the emergence of natural speech and results in a speech handicap which in turn results in a general linguistic retardation. This retardation responds to protracted educational treatment in a few cases, and intelligible speech and usable language are developed. More often, educational treatment fails to promote reasonably fluent speech and as this becomes apparent as pupils grow older, education becomes increasingly directed along non-oral lines, perhaps still verbally biased by silent reading, lip-reading and clandestine manual communication but possibly becoming more concerned with non-verbal areas such as artistic, technical, numerical and sporting activities.

In attempting to assess the role of communication skills in the education of the profoundly deaf it became obvious that the influence of residual hearing, intelligence and written language was considerable and that some partialling technique would be necessary. In consequence, a test battery suitable for factorial analysis was assembled with a substantial number of overlapping tests in order that factors might be identified more easily and that test results might be amply confirmed to offset the occasional unreliability of single tests which has been thought to occur in psychometric assessment of the deaf (Murphy: Ewing, 1957, p. 276).

A rotated factor description of the abilities of the profoundly deaf throws light upon many issues and controversies of practical education and suggests solutions based on a broad multifactor approach instead of the usual inappropriate consideration of an oversimplified model based on a few variables where irrelevant influences are not partialled out. Several such practical issues may be identified from the relationships set out in the following correlation and factor matrices and three of these current questions have been selected to demonstrate the practical applications of this kind of psychometric analysis. What is the influence of residual hearing upon educational attainment? What is the relationship of oral to manual communication skills? Is there a divergence between non-verbal intelligence and educational attainment? It may be noted that this approach deals with problems within the deaf population and is not concerned with the often pointless comparisons with hearing or partially hearing populations, although such comparisons are incidentally possible from the present data.

II.—Previous work.

No previous factorial analysis of the cognitive abilities of the profoundly deaf is known to the present writer. Presumably because standardised tests for the deaf must be administered individually, or in small groups of four at most, the majority of researchers have been content to register performance on a few tests. This invariably results in some factors being uncontrolled and occasionally the population-defining residual hearing variable has been left uncontrolled.

In some studies residual hearing is recorded but a low cut-off point between the partially and 'profoundly' deaf groups has resulted in including in the 'profoundly' deaf group many who have substantial usable residual hearing (and consequently good speech) together with those profoundly deaf who do not have sufficient residual hearing to exploit in speech training. For example, Gaskill (Ewing, 1957, p. 197) used a cut-off point between 'partially deaf' and 'deaf' at 60 decibels. More realistically, Wollman (1964), using an operational criterion, made his 'dividing line' at an average hearing loss of 75 decibels above normal threshold "to separate those children who hear unamplified speech albeit imperfectly and those with more severe impairments who, in the main, have little or no experience of unamplified speech." This cut-off level is equivalent to the 90 per cent. hearing loss cut-off level adopted elsewhere (Hood, 1949) and for purposes of comparison, it may be noted that, for each ear singly considered, an even loss of 60 decibels across the speech frequencies results in a hearing loss of 72 per cent. and an even loss of 75 decibels results in a 90 per cent. hearing loss. The conversion of pure-tone responses to percentage hearing loss for speech frequencies is made by means of tables constructed by Fowler and Sabine for the American Medical Association: the procedure is described by Hood (1949).

Another even more realistic educationally operational cut-off point would obviously be that which separated those who cannot perceive speech even with benefit of the standard powerful classroom amplification devices from those who have usable residual hearing which permits reception of amplified speech. A cut-off at the 75 decibels level could define a sample whose mean hearing loss was low enough to include a large majority with usable residual hearing and thus the mean hearing loss of a sample, together with its standard deviation should be stated.

Whenever partially deaf persons or those with usable residual hearing are included in a profoundly deaf sample, results are not applicable to the strictly defined profoundly deaf population which differs considerably in psychometric performance. The assumption of homogeneity even within the profoundly deaf group is questionable and it has been noted (Montgomery, 1966b, 1967) that the profoundly deaf group with 100 per cent. hearing loss in the better ear is notably poorer in speech than the profoundly deaf group as a whole :

	n	*Hearing Loss*	*Articulation Loss*
Total sample	82	95·1%	41·6%
Sub-group of sample	21	100·0%	51·2%

It may be further noted that the inclusion of substantial numbers of those with considerable residual hearing in test samples of the profoundly deaf, in addition to producing a misleading representation of the facts, has an important effect on teaching practice by lending specious support for those who maintain that all deaf children should be taught by the oral method (Ewing, 1951 ; Watson, 1967). Oral methods are not appropriate to the needs of the majority of the sub-group with 100 per cent. hearing loss, yet teachers in training are not instructed in any other teaching methods and consequently direct most of their efforts towards children with some effective residual hearing. It is easier to maintain the oral position when the non-oral group is spread thinly on the ground, outnumbered by the others and less clearly differentiated. Hence, the less strictly the deaf population is defined, the more will findings be compatible with the oralist claims.

By the so-called oral method, which is the official educational treatment throughout Britain, no finger spelling or signing is permitted as it is maintained that manual communication is detrimental to the development of the oral skills of speech and lip-reading. Yet, despite constant prohibition of manualism in school, no less than 71 per cent. of deaf children have been estimated (by teachers' ratings) to be able to communicate fluently with the adult deaf by manual methods as against the mere 7 per cent. who were estimated to be able to produce fairly fluent speech intelligible to the man-in-the-street (Montgomery, 1966b). The contention that there is a negative relationship between oral and manual communication skills has been questioned by Myklebust (1960), Birch and Stuckless (1964) and Montgomery (1966b). Birch and Stuckless (1964) further concluded that the influence of early manual communication, when present, tended to facilitate the development of other language attainments.

Simpson (1965), in a survey of the 1947 year group in English schools for the deaf, registered surprise at the low proportion of intelligent deaf children to gain admittance to the two selective schools for the deaf in Britain. In reply Askew (1965) presented a correlation matrix of the battery of selection test results which indicated substantial negative correlations between non-verbal performance I.Q. and verbal attainment tests and Montgomery (1966a) suggested that these negative relationships were due to selectivity effects.

III.—METHOD.

(*a*) *Population.*

Seventy deaf children from seven Scottish schools were assessed. These children are from the deafest section of the community, having a mean hearing loss for speech frequencies in the better ear of 96·1 per cent. (standard deviation 6·8 per cent.), one third of them, twelve boys and ten girls, registering a hearing loss of 100 per cent. All of this group were prelingually deaf, twenty-one congenitally deaf, ten congenitally deaf with deaf parents or siblings, sixteen deaf after meningitis, seven from other causes and sixteen whose deafness occurred from unknown causes.

This sample includes all profoundly deaf children born in Scotland in 1948 but all tests were not completed for two boys and two girls who left school before the recognised leaving age. Of the sixty-six who completed the test battery, thirty-five were boys and thirty-one girls. The age range of the group is from 15·6 to 16·2, with an average age of 15·11.

(*b*) *Instrumentation.*

Cognitive Tests.—Intelligence, attainment and perceptual tests are listed in Table 1 and numbered 4 to 22 for reference. Tests 4 to 12 are intelligence tests, 4 to 9 of the performance type and 10 to 12 of the paper and pencil type; 13 to 19 are language tests and 20 to 22 are numerical tests. Test 19 and 20 depend much upon perceptual speed rather than upon formal educational attainment. Tests 4 to 8 form part of the Snijders-Oomen Non-Verbal Scale standardised on European deaf children and tests 13, 14, 17, 18 and 20 have been specially devised for the profoundly deaf and are standardised on profoundly deaf Scottish children. Information about the availability of these tests is obtainable from the writer.

Ratings.—Teachers' estimates of each pupil's ability were obtained via head teachers and are listed in Table 1 (numbered 23 to 26 for reference). The use of concrete statements to describe ability levels was chosen to offset the tendency of teachers to employ different standards of assessment and the statements were deployed along a horizontal line with a sub-division of each statement into upper and lower ability groups, giving ten categories per rating, thus:

23. *Lip-reading.* A rating of oral comprehension under good lighting conditions using a personal hearing aid.

$\left.\begin{matrix}1\\2\end{matrix}\right\}$ No really useful grasp of what people say: recognises a few simple words only.

$\left.\begin{matrix}3\\4\end{matrix}\right\}$ Can recognise a few commonplace words and phrases when spoken deliberately.

$\left.\begin{matrix}5\\6\end{matrix}\right\}$ The ordinary deaf child in this respect: average understanding of the spoken word.

$\left.\begin{matrix}7\\8\end{matrix}\right\}$ Above average but falls short of understanding at the normal conversational rate.

$\left.\begin{matrix}9\\10\end{matrix}\right\}$ Can follow a normal conversation reasonably well.

24. *Voice Production.* A rating of speech intelligibility.

$\left.\begin{matrix}1\\2\end{matrix}\right\}$ Cannot produce recognisable words.

3 4	Very laboured speech which may be understood by those who have taught it. Spoken vocabulary very limited.
5 6	Can produce speech which is understood by those familiar with it at home and school.
7 8	Speech may be occasionally intelligible to others besides family and teachers.
9 10	Can produce fairly fluent speech, intelligible to the man-in-the-street.

25. *Written Composition.* A rating of written English.

1 2	Virtually illiterate : can write a few simple words only.
3 4	Can produce many written words and phrases but does not write in sentences.
5 6	Many errors of spelling and grammar but does try to write in sentences if left to own resources.
7 8	Writes in simple sentences with occasional grammatical slips.
9 10	Can produce good written work in grammatical sentences without assistance.

26. *Manual Communication.* A rating of finger spelling ability.

1 2	Has no knowledge of finger spelling.
3 4	Has some knowledge of finger alphabet but rarely uses it.
5 6	Average competence at understanding finger spelling.
7 8	Can communicate fluently with the adult deaf by using finger alphabet.
9 10	Almost total reliance on fingers for means of communication.

The distributions of teachers' ratings were not normal in some cases and a regrouping of categories was made. Lip-reading categories 5 and 6 were treated as a single category as were categories 5 and 6 of the Voice Production rating. Categories 4, 5, 6 and 7 in the Manual Communication rating were combined to offset the heavily skewed distribution.

Residual hearing was measured by reflecting the percentage hearing loss value for the better ear which was derived from pure-tone audiometric responses under clinical conditions converted into percentage values using the tables constructed for the American Medical Association.

(*c*) *Procedure.*

Tests were administered during the years 1963 to 1966 at the time of school-leaving usually in a pupil's last term at school. Most pupils were tested individually, even with group tests, but occasionally a maximum of four at a time were given a group test by the use of a language laboratory screening arrangement which was necessary to prevent distraction and the exchange of information by signing. Tests were carried out in all Scottish Schools for the Deaf, save for one session in the local Institute when the school was closed during an outbreak of typhoid.

Teachers' estimates were made independently and added to the battery to serve as validation criteria for the standardised tests and as corroborative aids to the identification of factors. All variables save the sex, age and aetiological categories were correlated to all others and the resultant product moment correlation coefficients assembled in the correlation matrix shown at Table 1. The correlation matrix was then submitted to a principal components analysis giving a general factor and 23 bipolar factors. Inspection showed clearly that the bipolar pattern is an inappropriate description of cognitive abilities which are more meaningfully adapted to a factor pattern which accords to the criterion of positive manifold. To meet this need for more psychologically meaningful factors a series of graphical rotations of the first eleven factors gave the solution shown in Table 2. This solution clearly distinguishes identifiable factors which make good psychological sense. It is not, of course, mathematically unique and further rotation to increase positive manifold and improve the factor structure was considered. But the communality value for each variable indicated that the approximations of succeeding rotations were accumulating and that further rotation would lead to solutions which, however psychologically elegant, would be unacceptably approximate.

IV.—RESULTS.

The main findings of this investigation are set out in Tables 1 and 2. Because, as has been noted, the factor solution given in Table 2 is not unique, the complete correlation matrix is reproduced in Table 1 to enable alternative rotated factor solutions to be computed if desired.

Children with 100 per cent. hearing loss in the better ear were significantly poorer in speech ability than the sample as a whole :

	n	*Hearing Loss*	*Articulation Loss*
Total sample	66	96·1%	38·8%
Sub-group of sample ..	22	100·0%	48·5%

A few sex differences were apparent : point biserial correlation coefficients indicated that girls were significantly poorer at the ·01 level in non-verbal reasoning as assessed by the S.O.N. Figure Analogies and Raven's Matrices tests and significantly better at the ·05 level in the Mill Hill Synonyms test and in Picture Vocabulary A.

V.—DISCUSSION OF RESULTS.

Inspection of the correlation matrix shows that no significant negative correlations obtain. It is of interest, however, that the slight negative relationship between the Block Design test 4 and residual hearing which is significant when deaf and hearing populations are compared, is still observable in the present population which is highly selected on the residual hearing variable.

The principal components analysis derived from the correlation matrix yielded a general factor and 23 bipolar factors. As they stand, these are psychologically meaningless configurations and, in the interests of economy, the factor matrix has not been reproduced here but is available from the writer on request. A rotated solution is presented in Table 2 where it may be noted :

Factor I is positively loaded on the three numerical tests. Factor II shows a relatively specific loading on the memory test 8. Factor III has a high loading on the rating of Manual Communication and some slight positive loading on the Mill Hill Vocabulary A (Definitions) test 15 and on both Picture Vocabulary

Tests 17 and 18. Factor IV has a high relatively specific loading on the McComisky test 9 with positive loadings on tests with a visual perception component 5, 19 and 20. Factor V has a high loading on the residual hearing test and substantial loadings on speech variables 14 and 24. Factor VI is a bipolar factor contrasting tests 5, 6 and 4, with some attainment tests. Factor VII is not easily identifiable as a meaningful factor but it has negative loadings on the two tests of perceptual speed 19 and 20. Factor VIII has high loadings on the two Lip-Reading variables 13 and 23. The relatively high loadings on other verbal and numerical tests probably indicate the importance of good general attainment for the development of ability in Lip-Reading. Factor IX contrasts tests 6 and 16 and is not easily identifiable. Factor X has high loadings on the intelligence tests and may be assumed to be a factor of non-verbal reasoning. The importance of this factor in the mathematics test 22 may be noted. Factor XI has high loadings on written vocabulary tests 15, 16, 17 and 18. The verbal content of part of mathematics test 22 accounts for its relatively high loading with this factor. Speech tests 14 and 24 and Lip-Reading test 13 obviously have a vocabulary component which accounts for their relatively high loading on this factor as has, less obviously, intelligence test 12 which uses a few words in some test items.

VI.—GENERAL DISCUSSION AND CONCLUSIONS.

From the relationships set out in the correlation and factor matrices it is possible to trace the interaction of many test variables and factors. For instance, the rather elusive concomitants of lip-reading (Simmons, 1959) may be observed in operation at the test or rating level or at the factor level. Similarly, the role of written vocabulary, number and perceptual speed in the educational abilities of the deaf may be separately assessed with regard to other abilities. The alleged divergence of concrete (performance test) intelligence and abstract (written test) intelligence may be examined as can the differences between tests which present all the information visually before the subject as against those which demand the recall of information or some ability to perceive seriation and to extrapolate.

But this is not the place to attempt to extract all the information from the results of this investigation and a return to the three current questions posed above ends this digression.

What is the influence of residual hearing upon educational attainment?

The correlation matrix reveals no significant relationship of hearing with any intelligence, attainment or communication variables save the two speech variables (r = ·52 and ·36). Thus, within this population with a narrow range of hearing loss, speech is still dependent to a substantial degree upon the possession of residual hearing. In the factor solution of Table 2, the relatively specific residual hearing factor with positive loadings on the speech variables has been taken out so that its influence upon other interrelationships has been removed.

Given a substantial relationship between speech and residual hearing, the relevance of the oral approach to the communication needs of those without usable residual hearing may be questioned. The present results confirm previous work which indicates that the mean articulation loss for the group with 100 per cent. hearing loss is well below the level required for the production of speech which is intelligible to listener or lip-reader. On the other hand, out of this same group with 100 per cent. hearing loss, teachers estimate that a majority can communicate fluently by means of the manual alphabet and all but four out of twenty-two can understand manual communication with at least 'average competence.' Is it not then reasonable to consider the possibility that the majority of those without usable residual hearing might be permitted to acquire an education via an approach which reinforces the oral instruction, which largely eludes them, with the manual method of communication which they largely understand already?

What is the relationship of oral to manual communication skills?

The alleged incompatibility between oral and manual communication skills is used as a justification for the prohibition of manual methods often by invoking the analogy of bad money driving out the good. Of course, there is a practical incompatibility at the classroom level where it is often expedient to prevent manual communication in order to concentrate upon speech training—just as in general language teaching it is natural to discourage the use of French conversation in a Latin lesson. From this it does not follow that oral and manual skills, or for that matter French and Latin skills, are negatively related and, in fact, some positive relationship might be expected on the basis of much similarity of linguistic content.

In the present study where the distorting effects of the residual hearing variable have been diminished no negative relationships between oral and manual skills are observable in the correlation matrix and teachers' assessments of manual communication are positively related to standardised tests 13 and 14, lip-reading and speech (r=·34, ·25).

The positive relationship between oral and manual skills was not maintained in the factor solution where the teachers' assessments of oral ability were given equal weight with the results of standardised tests. Another influence tending to produce unrelated speech and manual communication factors was the coarse class interval adopted to offset the heavily skewed distribution in the manual communication rating due to the large cluster of pupils in the 'fluent' categories 7 and 8. This coarse grouping probably depressed the communality of the manual communication variable resulting in the emergence of Factor III (manual communication) as a relatively specific factor but with positive loadings on the picture vocabulary tests. In contrast, the speech variables showed positive relationships with most of the language variables, high positive loadings on Factor V (hearing) and Factor VIII (lip-reading) and zero loadings on Factor III.

Is there a divergence between non-verbal intelligence and educational attainment?

The correlation matrix shows no significant negative relationships between performance tests and attainment variables. The correlation coefficients between the performance tests and language tests are, however, quite low. This tendency is even more apparent in Factor X (non-verbal reasoning) which has six zero loadings on language tests. Again, Factor XI (vocabulary) shows slight negative or zero loadings on all performance tests 4 to 9 as against positive loadings on the written intelligence tests, 10 to 12, and the bipolar Factor VI contrasts performance tests 4, 5 and 6, with attainment variables, including written composition and the two lip-reading variables. Thus, there is some basis for concern at the divergence of intelligence and verbal attainment (Askew, 1965) which suggests that there is a need for educational methods to exploit the abilities of the intelligent child who does not acquire language under present methods of instruction. That this is not merely a question of under-

TABLE 1

CORRELATION MATRIX OF TEST BATTERY AND TEACHERS' RATINGS.

	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
4. SON Block Design	1	35	55	65	33	34	67	54	57	27	17	29	15	22	26	27	37	57	60	26	24	29	09	*13*
5. SON Picture Memory ...		1	42	29	27	20	25	16	27	12	13	12	05	11	*03*	02	24	24	27	17	20	09	06	22
6. SON Picture Series			1	56	40	22	65	51	53	34	27	39	16	33	26	37	40	44	57	30	31	38	08	*02*
7. SON Figure Analogies....				1	45	45	74	69	67	38	30	40	17	30	32	42	50	56	64	22	29	38	30	01
8. SON Knox Cubes........					1	19	39	43	40	33	21	18	06	02	07	15	28	36	41	25	11	20	19	04
9. McComisky 3D Form Board						1	36	26	33	21	*01*	15	09	06	07	31	45	21	26	12	*02*	11	19	*02*
10. Matrices, 1938							1	68	67	42	35	45	17	33	35	44	51	61	72	28	28	38	34	07
11. APU Attributes								1	66	44	42	42	32	34	36	44	49	58	61	42	35	38	09	05
12. APU Abstractions									1	63	48	61	48	56	61	56	56	70	78	34	28	47	26	*04*
13. Lip-reading Test										1	75	75	57	77	84	80	71	46	72	75	74	85	34	08
14. Phoneme County (Speech)											1	74	49	69	68	55	49	56	59	57	80	69	25	52
15. Mill Hill Vocabulary A....												1	51	78	76	73	58	52	72	39	48	66	36	10
16. Mill Vocabulary Hill B....													1	60	61	52	35	29	45	35	40	49	15	03
17. Picture Vocabulary A														1	91	79	48	45	61	42	49	67	36	07
18. Picture Vocabulary B															1	83	50	49	67	49	50	68	40	03
19. Visual Word Discrimination																1	68	44	68	46	47	69	40	*06*
20. Number Perception																	1	52	64	56	41	52	35	07
21. Four Rules (Arithmetic) ..																		1	81	39	35	51	19	02
22. Vernon Maths.																			1	45	44	62	29	*02*
23. Lip-reading																				1	63	60	17	10
24. Voice Production																					1	67	10	36
25. Written Composition																						1	19	06
26. Manual Communication ..																							1	02
27. Residual Hearing																								1

NOTE.—Negative correlations are printed in italics. Decimal points are omitted.

TABLE 2

ORTHOGONAL FACTORS DERIVED FROM GRAPHICAL ROTATION OF PRINCIPAL COMPONENTS

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	h^2
TESTS :												
4. SON Block Design	14	*07*	*14*	09	*21*	42	30	29	29	63	06	1·08
5. SON Picture Memory	26	26	03	30	27	68	29	24	*17*	04	*14*	1·04
6. SON Picture Series	00	33	*13*	15	*02*	48	*05*	50	47	23	*09*	1·03
7. SON Figure Analogies............	02	13	05	*02*	06	18	24	27	42	63	00	0·89
8. SON Knox Cubes................	14	81	01	*72*	01	*02*	*04*	00	00	56	*14*	1·06
9. McComisky 3D Form Board	00	00	07	86	*05*	00	00	03	00	45	*04*	1·02
10. Matrices, 1938	02	*04*	15	06	10	33	*07*	33	48	65	12	1·06
11. APU Attributes	30	06	*16*	*07*	12	11	20	38	26	67	36	1·11
12. APU Abstractions	28	16	05	10	*04*	26	*08*	30	11	52	56	1·04
13. Lip-reading Test	18	15	12	18	07	*20*	04	78	16	08	43	1·02
14. Phoneme Count (Speech)	21	08	*02*	*06*	50	*01*	01	55	20	22	36	0·96
15. Mill Hill Vocabulary A	15	08	24	10	09	10	*12*	51	24	07	64	1·04
16. Mill Hill Vocabulary B	08	10	*16*	*05*	*01*	14	*09*	29	*37*	22	69	0·89
17. Picture Vocabulary A	04	00	32	01	06	11	00	36	17	*04*	74	0·92
18. Picture Vocabulary B	04	00	33	02	00	*06*	00	36	18	*03*	73	0·89
19. Visual Word Discrimination	04	01	26	28	*08*	*08*	*15*	60	30	10	57	1·07
20. Number Perception	36	*05*	14	42	17	00	*29*	65	29	28	18	1·18
21. Four Rules (Arithmetic)	47	*11*	01	17	23	*29*	10	54	*02*	00	24	0·84
22. Vernon Maths.	35	11	04	04	*03*	22	00	42	30	40	41	0·89
RATINGS :												
23. Lip-reading	27	08	*11*	06	11	*22*	03	90	00	13	*02*	1·07
24. Voice Production	20	*11*	00	*05*	39	*10*	22	46	16	06	54	0·96
25 Written Composition	11	15	*05*	10	04	*21*	23	61	32	01	48	0·99
26. Manual Communication	*01*	02	89	*01*	05	*03*	*02*	27	01	26	00	1·00
27. Residual Hearing	00	00	00	00	97	00	00	00	00	00	00	1·11

NOTE.—Values in italics are negative. Decimal point omitted.

attainment is shown by the positive correlation of the numerical tests with intelligence tests.

In summary, it may be concluded :

(1) Speech development is positively associated with the amount of residual hearing present even within the prelingually, profoundly deaf population.

(2) There is no significant negative relationship between any of the cognitive abilities assessed in this study. More particularly :

(*a*) There is no negative relationship between oral and manual communication skills.

(*b*) There is no negative relationship between non-verbal intelligence and verbal attainment.

(3) Some implications of these results are suggested as relevant to the oral approach to the education of deaf children which is the official method throughout Britain and currently the only method presented to teachers in training :

(*a*) The techniques for the development of speech in deaf children are notably unsuccessful with the majority of those who have no usable residual hearing.

(*b*) The alleged incompatibility of oral and manual methods used as a justification for the prohibition of the latter may be questioned.

(*c*) The low positive relationship between performance test results and verbal attainment implies a need for alternative educational methods to gear the intelligence of deaf children more effectively to academic achievement.

Acknowledgements.—The National Survey from which these results are sampled was commissioned by the Donaldson Trust, in 1963, and still continues. The writer is further indebted to the Applied Psychology Unit, University of Edinburgh, for the provision of psychometric material and the University Computer Unit for the provision of computer facilities, instruction and guidance.

VI.—References.

Askew, R. (1965). Grammar school entrance test results. *J. Nat. Coll. Teachers of the Deaf*, 63, 334.

Birch, J. W., and Stuckless, E. R. (1964). *The Relationship between Early Manual Communication and Later Achievement of the Deaf.* Co-operative Research Project 1769, U.S. Department or Health and Education Welfare.

Ewing, A. W. G. (1951). The oral way of living. *J. Nat. Coll. Teachers of the Deaf*, 49, 289.

Ewing, A. W. G. (Ed.) (1957). *Educational Guidance and the Deaf Child.* Manchester University Press.

Hood, H. B. (1949). A preliminary survey of some mental abilities of deaf children. *Brit. J. Educ. Psychol.*, 19, 210.

Kryter, K. D., Williams, C., and Green, D. M. (1962). Auditory acuity and the perception of speech. *J. Acoust. Soc. Am.*, 34, 1217.

Montgomery, G. W. G. (1966a). Difference in the interrelationships of non-verbal intelligence test scores, educational attainments and residual hearing in selected and unselected populations. *J. Nat. Coll. Teachers of the Deaf*, 64, 29.

Montgomery, G. W. G. (1966b). The relationship of oral skills to manual communication in profoundly deaf adolescents. *American Annals of the Deaf*, 111, 557.

Montgomery, G. W. G. (1967). Analysis of pure-tone audiometric responses in relation to speech development in the profoundly deaf. *J. Acoust. Soc. Am.*, 41, 53.

Myklebust, H. R. (1960). *The Psychology of Deafness.* New York : Grune and Stratton.

Oléron, P. (1950). A study of the intelligence of the deaf. *American Annals of the Deaf*. 95, 179.

SIMONS, AUDREY A. (1959). Factors related to lip reading. *J. Speech Hear. Res.*, 2, 340.

SIMPSON, E. (1965). *The Health of the School Child*, 1962, *and* 1963. (Chapt. 9). H.M.S.O.

TAYLOR, I. G. (1967). Disorders of communication in deaf and hearing impaired children. *B.P.S. Bulletin*, 20, 34.

WATSON, T. J. (1967). *The Education of Hearing Handicapped Children*. London: University of London Press.

WOLLMAN, D. C. (1964). The attainments in English and Arithmetic of secondary school pupils with impaired hearing. *Brit. J. Educ. Psychol.*, 34, 268.

(*Manuscript received* 9*th February*, 1967)

TEACHERS' ATTITUDES TO EDUCATION

By R. A. C. OLIVER and H. J. BUTCHER
(*Department of Education, University of Manchester*)

SUMMARY. Three scales measuring attitudes to education were administered to a representative sample of 300 teachers. The results were analysed by sex, age, political party, religious affiliation and type of school. Tests of significance were applied, both to the relation between these biographical categories and to differences in attitude between the groups of teachers.

I.—INTRODUCTION.

THE construction of three attitude scales designed to measure naturalism in education, radicalism in education and tender-mindedness in education, and the selection of a representative sample of 300 teachers as respondents were described in an earlier paper (Oliver and Butcher, 1962). The use of these scales with student teachers has been reported elsewhere (Kissack, 1956; Steele, 1958; Thompson, 1957; Butcher, 1965). The present paper describes their operational use to assess the attitudes of practising teachers.

Information about the teachers' sex, age, type of school, political party and religious affiliation was obtained, and these five types of information formed the principal categories by which scores were analysed. Since the sample approximately reflected in miniature the distribution of all teachers in England and Wales in 1955, by sex, age and type of school, the relation between the categories is of some interest, and is briefly reported in the next section.

II.—RELATIONS BETWEEN THE BIOGRAPHICAL, POLITICAL AND RELIGIOUS CATEGORIES.

In the subsequent tables, the following classifications of biographical data were made:

(1) *Age groups.* Only nine of the 300 teachers in the sample were aged over 60. This small group was accordingly combined with the 50–59 age-group to make a 50 plus age-group containing sixty-nine teachers.

(2) *Type of school.* All-age and unreorganised schools were grouped with primary schools. Thirty-nine teachers were teaching in institutions of further education or in schools other than those in the three main categories (primary, secondary modern, secondary grammar). These thirty-nine were grouped together as one category, 'other types of school,' for tabulation and analysis. These thirty-nine teachers were distributed as follows: secondary technical schools, nine; special schools, seven; comprehensive schools, one; further education, nineteen; unspecified 'other', three.

(3) *Political party.* One Communist teacher was included in the Labour group. Of the forty-nine teachers classified as supporting no political party, forty-six stated this specifically, the remaining three left a blank space.

(4) *Religious affiliation.* The fifty-six Nonconformist teachers were distributed as follows: Methodists, forty-two; Congregationalists, twelve; Presbyterians, one; unspecified Nonconformist, one.

The twelve teachers placed in the 'miscellaneous' category in respect of religious affiliation were distributed thus: Jewish, three; Christian Scientist, three; Unitarian, two; Humanist, two; undenominational, two. The category 'no religious affiliation' includes one teacher who left the space blank.

TABLE 1

INTER-RELATIONS OF BIOGRAPHICAL DATA.

Comparison	Value of chi-square	d.f.	Level of significance of association
Sex and age	7·94	3	·05
Sex and type of school	42·36	3	·01
Sex and political party	12·49	3	·01
Sex and religious affiliation	5·78	4	n.s.
Age and type of school	6·54	9	n.s.
Age and political party	26·51	9	·01
Age and religious affiliation	21·77	12	·05
Type of school and political party	14·87	9	n.s.
Type of school and religious affiliation	18·35	9	·05
Political party and religious affiliation	74·40	9	·01

The significant association between age and sex in Table 1 arose from a preponderance of women in the highest age-group. There are markedly ' too many ' men in the 30–39 age-group and ' too few ' in the Over 50 age-group. The high value of chi-square in the case of sex and type of school was, as might be expected, due to the high proportion of women teachers in primary schools. There is a correspondingly high proportion of men in the other three categories, but the difference is small in the case of secondary grammar schools.

As to the relation between sex and political party, a disproportionately large number of men supported the Labour party, and to a lesser degree men also predominated in the ' no party ' category. Women, by contrast, predominated in the Conservative group. The proportions supporting the Liberal party corresponded closely to the ' expected ' distribution. Analysis in detail of the association between age and political party reveals some interesting results. The largest contribution to the value of chi-square was made, not surprisingly, by the relatively small number of teachers over the age of 50 supporting the Labour party, and the second largest by the relatively large number supporting the Conservative party. The same trend held, though to a much lesser extent, in the *youngest* (under 30) group. It was the two middle age-groups which, relatively speaking, furnished most support for the Labour party. The Liberal party tended to find relatively more supporters in the two highest age-groups, but particularly in the 40–49 group. This latter group also supplied more than its share of those who deny support of any political party.

Of those teachers who professed no religious affiliation a high proportion were aged between 30 and 40. Similarly, in this age-group relatively few were members of the Church of England; so few that in all other age-groups the number of Church of England members exceeded expectation. In the other three religious groups (Roman Catholic, Nonconformist and Miscellaneous) no marked association with age is apparent. The comparison of religious affiliation and the type of school in which the teacher was employed produced a significant value of chi-square, but the nature of the relationship is rather less clear than in most of those so far reported. Teachers in ' other types of school ' and to a lesser degree in secondary grammar schools, tended to fall in the ' no religious affiliation ' category. The secondary modern school group contained more than its share of Nonconformists and Miscellaneous. Teaching in primary schools appears to have been associated with being a member of the

Church of England, and to a lesser extent with being a member of the Roman Catholic Church. In the case of religious affiliation and political party, there is a very marked association between ' Labour ' and ' no religious affiliation,' and also between ' Liberal ' and ' Nonconformist ' and ' miscellaneous.'

III.—THE ANALYSIS OF TEACHERS' ATTITUDE SCORES BY SEX, AGE, TYPE OF SCHOOL, RELIGIOUS AFFILIATION AND POLITICAL PARTY.

Means and standard deviations of scores on naturalism in education, radicalism in education and tender-mindedness in education are shown below in Table 2. High scores represent naturalism, radicalism and tender-mindedness respectively.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF ATTITUDE SCORES IN RELATION TO AGE, SEX AND OTHER VARIABLES.

	Naturalism		Radicalism		Tender-mindedness		N
	M	S.D.	M	S.D.	M	S.D.	
Whole sample	74·87	9·68	112·63	15·95	128·47	16·16	300
Men	74·48	9·52	112·12	17·12	127·67	16·18	127
Women	75·15	9·79	113·01	15·03	129·05	16·13	173
Under 30	76·10	8·42	114·24	16·08	128·25	13·39	61
30–39	75·75	11·87	113·06	17·90	132·44	17·73	80
40–49	74·44	9·32	112·83	14·99	129·66	15·74	90
Over 50	73·30	7·39	110·62	14·29	122·36	15·26	69
Teachers in :							
Primary schools ..	75·88	10·11	112·49	16·58	127·35	15·84	173
Sec. Mod. schools	74·69	8·70	114·87	13·56	128·48	14·08	54
Sec. Gram. schools	69·79	8·51	107·88	17·25	136·12	14·87	34
Other schools	75·08	8·52	114·31	16·01	126·74	17·47	39
R.C.	71·33	8·06	106·17	14·50	124·00	14·48	30
C. of E.	73·75	8·66	111·52	15·84	126·05	16·24	149
N.C.	74·50	8·66	110·04	13·59	126·07	14·13	56
Miscellaneous	77·00	7·38	119·08	15·58	136·17	18·75	12
None	79·92	12·42	120·68	16·02	136·72	16·58	53
Conservative	73·17	9·72	107·75	14·44	124·85	16·17	146
Liberal	73·48	9·89	110·98	14·68	127·45	10·92	44
Labour	78·48	8·84	123·23	15·67	136·18	15·48	61
None	76·98	8·75	115·47	14·68	130·57	17·17	49

Tests of significance were applied to the differences between groups on all three scales. The results of these tests are summarised in Table 3. Where overall significant differences were found, by analysis of variance, these were explored in more detail by t-tests between pairs of groups, with the following results.

On the scale of naturalism, teachers in secondary grammar schools were significantly less naturalist than all other teachers. There were no significant differences between other types of school. Those teachers who professed no

TABLE 3

TESTS OF SIGNIFICANCE OF MEAN DIFFERENCES IN ATTITUDE BETWEEN GROUPS OF TEACHERS.

	Naturalism	Radicalism	Tender-mindedness
Sex	n.s.	n.s.	n.s.
Age	n.s.	n.s.	·01
Type of school	·05	n.s.	·05
Religious affiliation	·01	·01	·01
Political party	·01	·01	·01

religious affiliation were significantly more naturalist than any of the religious groups, but there were no significant differences between these latter. Conservative teachers were significantly less naturalist than Labour teachers, but there was no significant difference between Conservatives and Liberals.

On the scale of radicalism, those teachers who professed no religious affiliation were significantly more radical than teachers belonging to any of the three main religious groups (Roman Catholic, Church of England, Noncomformist), but there were no significant differences between these religious groups. Labour teachers differed significantly in respect of greater radicalism from Conservatives, from Liberals, and from teachers of ' no party ' ; the only other significant difference was that Conservative teachers were less radical than those of ' no party.'

On the scale of tender-mindedness, teachers over 50 were significantly more tough-minded than those in all other age groups, but no other age differences were significant. Teachers in grammar schools were significantly more tender-minded than all other teachers, with no such differences among the remaining groups. Teachers with no religious affiliation were significantly more tender-minded than all the main religious groups, with no such differences among the latter. Labour teachers were more tender-minded than either Liberals or Conservatives, with no significant difference between the latter groups.

IV.—DISCUSSION.

To consider first the inter-relations among the biographical data, statistically significant associations were found in seven of the ten comparisons, the three exceptions being sex and religious affiliation, age and type of school, political party and type of school. The results of the seven comparisons in which significant associations were found appeared to be generally in line with expectation. The strongest degree of association was found between political party and religious affiliation (a coefficient of contingency of ·46) ; support of the Labour party was associated with no religious affiliation and support of the Conservative party with membership of the Church of England. In this respect and in some details revealed by the contingency table, the teachers who were the subjects in the present research appeared to correspond fairly closely with the whole population of Great Britain, as described by Eysenck (1954, page 21) in reviewing some of the results obtained by the British Institute of Public Opinion. The same is not true, however, of the relation between sex and political party. The opinion poll results show that there was only a slight tendency in the whole population for women to favour the Conservative party and men the Labour party, whereas in the sample of 300 teachers the association was very much more marked.

The association between age and political party was, as has been stated, found to be highly significant, but not quite in the manner expected. It might have been supposed that the proportion of teachers supporting the Conservative party would increase fairly steadily with age, and that there would be a corresponding decrease in the proportion of Labour supporters. A clear, almost linear, trend of this kind was found by the British Institute of Public Opinion in the whole population of the country. In the sample of teachers, on the other hand, the proportion supporting the Conservative party showed a marked drop after the age of 30 and rose again only after the age of 50. The pattern of Labour party support was similar, but in reverse, indicating relatively more support in the middle age-groups (between 30 and 50). Support of the Liberal party followed a trend somewhat similar to that of the Labour party, but with a greater concentration in the upper two age groups (i.e., among people of over 40).

Results of the analysis of attitude scores by sex, age, type of school, religious affiliation and political party were summarised in Table 2. Disregarding for the moment any tests of significance, one can see at once certain consistencies and regularities in this table of results. Firstly, with very few exceptions, the separate groups in each biographical category are put in the same order by all three scales. Thus, women score slightly higher on the average than men on naturalism, radicalism and tender-mindedness. Similarly, of the political groups, the Labour party supporters scored highest on all three scales, the Conservatives lowest, and the two other groups (Liberal and 'no political party') are also placed in the same relative position on each scale. There are slight exceptions to this consistency of pattern in the breakdowns by age, type of school and religious affiliation, but even here the regularity is striking. The older the group of teachers, the more idealist, conservative and tough-minded as described by their mean score they were, with only one exception; the youngest (under 30) group has a lower mean score on the T scale than either of the middle age-groups; and this finding may be connected with the similar pattern reported in the analysis of the biographical data, in which it was found that the youngest age-group was relatively more Conservative in politics than the two middle age-groups. In the breakdown by type of school, the salient feature is the difference in mean score between teachers in grammar schools and all other teachers. It is apparent from Table 2 that the former obtained a more idealist, more conservative, but more tender-minded score than any of the other groups, between which the differences are quite small. It is possible that this high score of grammar school teachers on the T scale reflects the 'cultural' *versus* 'technical' aspect of the scale.

Finally, the analysis by religious affiliation again presents a similar regularity. Teachers with no religious affiliation obtained the highest mean score on all three scales, closely followed by those in the 'miscellaneous' category (e.g., Jewish, Christian Scientist, Unitarian). Roman Catholic teachers had the lowest mean score on each scale. Church of England and Nonconformist teachers obtained closely similar mean scores on all scales, above the Roman Catholics, but below the 'miscellaneous' group.

It appears rather surprising that the three scales should have ordered the groups in all these categories so similarly, particularly when one considers that the correlation between the R and T scales was only $+\cdot 31$. It was not within the scope of this research to repeat the analysis with the purer measures provided by the factor scales which have also been developed, but this would clearly be an interesting subject for further research.

Tests of significance showed that not all the trends just described were statistically significant. Sex differences were not significant on any scale; nor were age differences on the N and R scales. The oldest (over 50) group was, however, significantly more tough-minded than any other group; the rather more tough-minded mean score of the youngest did not differ significantly from the mean scores of the two middle age-groups. Tests of significance applied to the differences between types of school showed that teachers in grammar schools were significantly more idealist and more tender-minded than all the other groups. Significant differences on all scales were found between the different religious and the different political groups. More detailed tests showed that the main dividing line of opinions lay between those with no religious affiliation on the one hand, and all the religious groups on the other. In the detailed comparison of political groups, it was interesting to find that the Liberals differed significantly from the Labour group on all three scales, but from the Conservative group on none.

Most of these findings are in general accord with such earlier findings as exist in this field. To take first the research with the most closely similar subject-matter, the findings of Peterson (1933) on teachers' opinions in the U.S.A. were in general confirmed; it is admittedly difficult to make a close comparison owing to the different terms of reference of the two researches, but at the least no obviously opposed conclusions emerge. In particular, Peterson found a correlation between the active practice of religion and a 'traditional' or idealist approach to educational problems. Some findings of Cook, Leeds and Callis (1951), with the Minnesota Teacher Attitude Inventory, were also paralleled in the present research, viz., the more 'naturalist' attitude of student teachers as compared with that of practising teachers, the smaller spread of scores among the former than among the latter, and the more 'idealist' attitude of the American teachers corresponding to grammar school teachers. Shapiro's (1952) finding (based on a sample of parents) that "subjects who voted Labour would tend to be tolerant in their attitude to the free expression of children's desires, while those subjects who vote Conservative would tend to be intolerant of such expression" is shown by the present research to apply also to teachers.

The common findings so far discussed have not included correlations between attitude and sex or age, since these appear to have been generally small, or at best ill-established. In the field of educational attitudes, Kissack (1956) found women to be significantly more opposed than men to corporal punishment, but apart from this, few striking sex differences appear to have been reported. Peterson (1933) found younger lecturers in American training colleges to be more 'progressive' than older ones, but Koch, *et al.* (1934) found no significant age difference. In terms of general social attitudes, Ferguson found older men significantly less 'humanitarian,' but found no change with age among women; Eysenck reported no significant sex or age differences in radicalism, but a significant sex difference in favour of women on tender-mindedness. It may, therefore, be concluded that differences in attitude dependent on age and sex are rather less marked than some of the other differences discussed (e.g., in politics and religion). This conclusion is not disturbed by the present research, in which the only such significant trend found with the main sample was an association between age and toughmindedness in education.

V.—REFERENCES.

BUTCHER, H. J. (1959). *The Opinions of Teachers and Student Teachers About Education* Unpublished Ph.D. thesis. University of Manchester Library.

BUTCHER, H. J. (1965). The attitudes of student teachers to education. *Brit. J. Soc. Clin. Psychol.*, 4, 17-24.

COOK, W. W., LEEDS, C. H., and CALLIS, R. (1951). *Manual of the Minnesota Teacher Attitude Inventory.* New York: Psychological Corporation.

EYSENCK, H. J. (1954). *The Psychology of Politics.* London: Routledge.

KISSACK, M. (1956). *The Attitudes of Training College Students Towards Corporal Punishment.* M.Ed. thesis. University of Manchester Library.

KOCH, H. L., KENTLER, M., DYSART, B., and STREIT, H. (1934). A scale for measuring attitude toward the question of children's freedom. *Child Development*, 5, 253-266.

MCINTYRE, D., and MORRISON, A. (1967). The educational opinions of teachers in training. *Brit. J. Soc. Clin. Psychol.* 6, 32-37.

MCINTYRE, D., MORRISON, A., and SUTHERLAND, J. (1966). Social and educational variables relating to teachers' assessments on primary school pupils. *Brit. J. Educ. Psychol.*, 36, 272-279.

MORRISON, A., and MCINTYRE, D. (1967). Changes in opinions about education during the first year of teaching. *Brit. J. Soc. Clin. Psychol.* 6, 161-163.

OLIVER, R. A. C. (1953). Attitudes to Education. *Brit. J. Educ. Studies*, 2, 31-41.

OLIVER, R. A. C. and BUTCHER, H. J. (1962). Teachers' attitudes to education. The structure of educational attitudes. *Brit. J. Soc. Clin. Psychol.*, 1, 56-59.

PETERSON, R. E. (1933). Philosophies of education current in the preparation of teachers in the United States. *Publ. Teach. Coll.*, Columbia University.

SHAPIRO, M. B. (1952). Some correlates of opinions on the upbringing of children. *Brit. J. Psychol.*, 43, 141-149.

STEELE, P. M. (1958). *Changes in Attitude Amongst Training College Students Towards Education in Junior Schools.* M.Ed. thesis. University of Manchester Library.

THOMPSON, J. W. (1957). *A Factorial Study of the Values and Attitudes of Graduate Students in Training.* M.Ed. thesis, University of Manchester Library.

(*Manuscript received* 3rd *March*, 1967)

VARIATIONS IN PERFORMANCE IN A UNIVERSITY ENTRANCE TEST IN ENGLISH

BY R. A. C. OLIVER AND D. G. LEWIS
(Department of Education, University of Manchester)

SUMMARY. The performance of candidates in a university entrance test in English was analysed in terms of arts-science choice in G.C.E. Advanced level subjects, type of school and size of sixth form. Generally, arts candidates did better than science candidates, and girls did better than boys. A significant second-order interaction between school type, sixth-form size and sex was revealed.

I.—INTRODUCTION.

THE Joint Matriculation Board (J.M.B.) accepts as a university entrance qualification a pass in an approved test in English, and itself conducts such a test twice a year. In a report on the Board's Experimental Test in English held in 1962, Oliver reported differences in marks between schools, between boys and girls and between ' Arts ' and ' Science ' candidates (Oliver, 1963). Candidates may, however, be cross-classified by school, sex and subject specialisation, and the data previously presented did not reveal the relationship between marks and the possible categories of cross-classification. This the present paper does in respect of the Board's Test in English held in March, 1965. The paper is offered as presenting results of interest to teachers and university selectors. It also illustrates how sampling at two levels (schools and pupils) can be incorporated into an experimental design.

II.—THE TEST IN ENGLISH.

The Test in English is a qualifying examination : a candidate's result is effectively pass or fail, though several grades of pass and fail are reported for information. The test is not intended to provide a pattern for English teaching in Sixth Forms, or to favour one school of thought about English teaching rather than another. It is a test intended to provide a qualification for admission to a university as an undergraduate in any Faculty—Arts, Science or other. In 1965, 87·3 per cent. of the candidates who took the test at one of the two opportunities in that year passed. A test serving such a purpose will be required to discriminate among candidates in the lower range of ability in English, rather than to reveal in full the brilliance of the ablest candidates.

The test is a three-hour one. In March, 1965, there were two questions, both of which candidates had to answer. Question 1 consisted of an objective test of reading comprehension, together with a short essay requiring ' personal ' writing. Question 2 may be described as testing knowledge of English usage, together with a composition requiring descriptive and personal writing.* The maximum mark obtainable in the paper was 75, the mean mark 44·75, the standard deviation of the marks 7·93. The pass mark was 39 (52 per cent. of the maximum). 78·5 per cent. of the candidates in March, 1965, passed.

---

* Copies may be obtained from John Sherratt and Son (Publishers), Park Road, Timperley, Altrincham, Cheshire.

III.—The schools and candidates.

The marks analyzed were those of a sample of the 19,818 candidates entered at 868 schools and centres.

Schools were classified by *type* as direct grant schools or maintained schools. Other types of school, e.g., independent schools and aided schools, were not sampled.

Schools were also classified by *size*, on the basis of the number of candidates entered for the test. The schools were placed in rank order in respect of the number of candidates entered. Schools in the top third of the rank order were categorised as large, the remainder as small.

Schools were also classified by *sex* as boys' schools or girls' schools. Mixed schools were not sampled.

Candidates were classified as 'Arts' or 'Science.' Those taking General Certificate of Education (G.C.E.) examinations at the Advanced level in 1965 or earlier, in 'Science' subjects (including Mathematics and Geology) were classified as 'Science' candidates, all others as 'Arts' candidates. This classification will be referred to as *choice* of specialisation.

As schools were classified by type (two categories), size (two categories) and sex (two categories), there are eight cross-classifications in all. Within each of these cross-classifications four schools were selected at random, and within each of these schools six 'Arts' candidates and six 'Science' candidates were selected at random. The analysis is, therefore, based on the marks of 8×4×12 candidates, a total of 384.

IV.—Statistical analysis.

(i) *Analysis of Variance.*

The complete analysis must distinguish, in the first place, variations arising from school type, size and sex and as three main effects, together with the three first-order interactions and the second-order interaction between these effects. It must distinguish also the different sources of variation within each of the school type by size by sex cross-classifications, namely those of (arts-science) choice, school and the interaction between choice and school. The analysis is thus complex, involving both the 'crossing' and 'nesting' of variables, and the appropriate model is set out in full in the Appendix. The results are shown in Table 1. The significance of each of the first seven effects—all those *between* the various school types by size by sex cross-classifications—is tested by the mean square for schools within school type by size by sex, while for the effects *within* school type by size by sex that of both schools and schools × choices interaction is tested by the mean square for pupils, and that of choices is tested by the mean square for the schools × choices interaction (see the mean square expectations set out in the Appendix). The F ratios shown in the last column of the table have been obtained on this basis. We see that statistically significant effects exist for sex, i.e. for boys' and girls' schools, for the second-order interaction school type × size × sex, and—turning to the effects within school type by sex by size—for both choices and schools.

Since the schools × choices (S × C) interaction is not significant, it would be sensible to combine this source of variation with that of choices (C), i.e., to combine lines 9 and 10 of Table 1, in an alternative analysis shown in Table 2. The sources of variation within school type by size by sex would then be seen to correspond exactly with the sampling procedure, namely that of pupils within (arts-science) choices, and choices within schools.

TABLE 1

ANALYSIS OF VARIANCE.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
School Type (T)	228·1	1	228·1	3·39
School Size (Si)	147·5	1	147·5	2·17
Sex (Se)	677·3	1	677·3	10·07**
T×Si	2·1	1	2·1	<1
T×Se	4·2	1	4·2	<1
Si×Se	8·8	1	8·8	<1
T×Si×Se	468·2	1	468·2	6·96*
Schools (S), within T×Si×Se	1,614·0	24	67·25	1·65*
Choices (C), within T×Si×Se	1,780·0	8	222·50	5·32**
S×C, within T×Si×Se	1,003·8	24	41·82	1·03
Pupils, within S×C	13,049·0	320	40·78	
Total	18,983	383		

** Indicates significance at the 1 per cent. level.
 * Indicates significance at the 5 per cent. level.

TABLE 2

ALTERNATIVE ANALYSIS OF THE VARIATION WITHIN SCHOOL TYPE BY SIZE BY SEX SHEWN IN TABLE 1.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Schools (S), within T×Si×Se	1,614·0	24	67·29	1·65*
Choices (C), within S	2,783·8	32	86·99	2·13**
Pupils, within C	13,049·0	320	40·78	

** Indicates significance at the 1 per cent. level.
 * Indicates significance at the 5 per cent. level.

(ii) *Differences within the School Classifications.*

The mean scores of all the arts and science pupils within each of the eight school classifications are shown in Table 3. The difference, arts pupils' mean score minus science pupils' mean score, is also shown. We see that in all cases the arts pupils have a superior test performance, this superiority being most marked for the small maintained grammar boys' schools and the large direct-grant girls' schools. Overall, the superiority of the arts pupils has been shown to be significant at the 1 per cent. level (Table 1).

There are, of course, appreciable differences between schools, even within each of the classifications. Table 4 shows the difference in mean score (arts pupils' mean score minus science pupils' mean score) for each of the separate schools. The differences between these mean score differences, within each of the classifications—i.e., the schools × choices interactions—are not, however, significant at the 5 per cent. level (Table 1). We see that in only four of the thirty-two schools is the mean score of the science pupils the greater. There is little interest in the combined differences between schools, i.e., differences between the mean score of the arts plus science pupils combined, shown to be significant at the 5 per cent. level. With random samples of schools, significant school differences are to be expected.

TABLE 3

MEAN SCORES OF ARTS AND SCIENCE PUPILS.

		Large Schools		Small Schools	
		Direct Grant	Maintained	Direct Grant	Maintained
Boys	Arts	48·54	46·79	46·17	48·37
	Science	46·62	41·58	43·00	42·25
	Difference (Arts-Science)	1·92	5·21	3·17	6·12
Girls	Arts	51·46	49·87	51·54	47·17
	Science	45·62	48·42	47·17	43·33
	Difference (Arts-Science)....	5·84	1·45	4·37	3·84

TABLE 4

DIFFERENCES (ARTS-SCIENCE) IN MEAN SCORE FOR THE SEPARATE SCHOOLS.

	Large Schools		Small Schools	
	Direct Grant	Maintained	Direct Grant	Maintained
Boys	2·83 7·00 —3·30 1·17	5·00 7·67 6·50 1·67	4·83 4·50 6·50 —3·17	—1·00 8·83 12·83 3·83
Girls	6·00 2·33 6·67 8·33	3·67 3·50 —4·50 3·17	4·67 7·33 1·00 4·50	6·17 2·17 3·50 3·50

(iii) *Differences between School Classifications.*

The mean scores obtained by all forty-eight pupils in each of the eight classifications of schools is shown in Table 5. We see that whereas girls do better than boys in three of the four type by size classifications—and markedly so in the large maintained grammar and small direct-grant schools—it is the boys who do better (if only very slightly) in the small maintained grammar schools. It is this difference in pattern which has been shown by the second-order interaction school type × size × sex of Table 1, to be significant at the 5 per cent. level.

TABLE 5

MEAN SCORES OF ALL PUPILS IN EACH OF THE CROSS-CLASSIFICATIONS OF SCHOOLS.

	Large Schools		Small Schools	
	Direct Grant	Maintained	Direct Grant	Maintained
Boys	47·58	44·19	44·58	45·31
Girls	48·54	49·15	49·35	45·25

Combining the scores for school type, size and sex, in turn, gives the mean scores shown in Table 6. No interactive effects are now suggested, so confirming the analysis of Table 1, the overall superiority of the direct grant schools over the maintained grammar schools, of girls over boys, and of large size over small being apparent. The mean scores for the three main effects are shown in Table 7. Only one of these differences, that for girls minus boys, has been shown to be statistically significant, and at the 1 per cent. level.

TABLE 6

MEAN SCORES FOR THE TYPE BY SIZE, TYPE BY SEX AND SIZE BY SEX CROSS-CLASSIFICATIONS OF SCHOOLS.

(i)

	Direct Grant	Maintained
Boys	46·08	44·75
Girls	48·95	47·20

(ii)

	Direct Grant	Maintained
Large	48·06	46·67
Small	46·97	45·28

(iii)

	Large	Small
Boys	45·89	44·95
Girls	48·84	47·30

TABLE 7

MEAN SCORES FOR DIFFERENT CATEGORIES OF SCHOOL TYPE, SIZE AND SEX.

(i)

Direct Grant	47·52
Maintained Grammar....	45·97

(ii)

Boys	45·42
Girls	48·07

(iii)

Large Schools	47·36
Small Schools	46·13

V.—CONCLUSIONS.

The main conclusions may be stated as follows :

(1) Within all eight cross-classifications of schools (type by size by sex) the performance of the arts pupils was superior to that of the science pupils, the overall extent of this superiority being significant at the 1 per cent. level.

(2) While within each of the cross-classifications of schools the arts-science differential itself varied from school to school, nevertheless, the schools and subject choice interaction was not significant at the 5 per cent. level.

(3) The extent of the arts pupils' superiority was found to differ considerably over the different cross-classifications of schools, being greatest for the small maintained grammar boys' schools and the large direct grant girls' schools (mean differences 6·12 and 5·84, respectively) and least for the large direct grant boys' schools and the large maintained grammar girls' schools (mean differences 1·92 and 1·45, respectively).

(4) The second-order interaction, school type × size × sex, was found to be significant at the 5 per cent. level. The performance of the girls was superior to that of the boys in three of the four school type by size cross-classifications (and markedly so in the large maintained grammar schools and small direct grant schools), while in the small maintained grammar schools the difference (in favour of the boys) was small and statistically insignificant.

(5) Overall, the superiority in the test performance of the girls was significant at the 1 per cent. level.

VI.—DISCUSSION.

This paper analyses the performance in a test of English of sixth-form candidates grouped in four different ways, those of school type, sixth-form size, sex of candidate and candidate's arts-science choice. The prime interest was that of investigating the effect of sex and of choice of subjects upon test performance. It was considered, however, that school type and sixth-form size might well have important interactive effects.

One point concerning the design may be noted. The four groupings cannot all appear as main effects in a typical four-way analysis of variance. This is because of the requirement of random groups. The pupils selected were random samples of arts and science candidates within each school, but even though the schools were randomly selected from each cross-classification of school type, size and sex, the pupils do not comprise a random sample of the whole (i.e., of each cross-classification). Random samples from sub-populations, even of randomly selected sub-populations, do not provide in combination a random sample of the population. Sampling at two levels, schools and pupils, has to be incorporated into the design.

The results indicate a general superiority in test performance of the arts candidates. This, in view of the greater role language assumes in their sixth form work, is not unexpected, and a similar test of 'numeracy' would probably reveal a corresponding superiority on the part of the science candidates. Almost equally important, however, is the lack of uniformity in this superiority over the different cross-classification of schools. If an improvement in test performance is to be achieved by teaching rather than by changes in school organisation or in the distribution of pupils between arts and science courses, then the greatest scope for improvement at present lies with the boys taking science courses in maintained grammar schools.

The lack of uniformity in the arts-science differences also has important consequences in any proposal for a relaxation of the English requirement for science candidates. The backwash effect that the implementation of such a proposal would have on the schools is speculative, and the scope of this study is insufficient to provide any evidence about such an effect. What is clear, however, is that if such a relaxation were applied, account would have to be taken of the interactive effects of school type, size and sex.

Of the other differences, that of sex–showing a general superiority in the test performance of girls—is the most pronounced. This need not arise from an intrinsic sex difference since in the small maintained grammar schools the general level in the performance of boys and girls is about equal. The pattern of results suggests that greater provision could effectively be made for the English studies of boys in both the large maintained grammar schools and the small direct grant schools.

VII.—REFERENCE.

OLIVER, R. A. C. (1963). *An Experimental Test in English.* Manchester : Joint Matricula-Board, Occasional Publication 13.

VIII.—APPENDIX.

THE MODEL UNDERLYING THE ANALYSIS OF VARIANCE.

With A, B, C, D and E referring to school type, size, sex, school and choice, respectively, the score x_{ijklmn} of the n^{th} pupil in the m^{th} choice in the l^{th} school in the k^{th} sex, j^{th} school size and i^{th} school type is made up of twelve independent components as follows :

$$x_{ijklmn}=M+A_i+B_j+C_k+(AB)_{ij}+(AC)_{ik}+(BC)_{jk}$$
$$+(ABC)_{ijk}+D_{ijkl}+E_{ijkm}+(DE)_{ijklm}+e_{ijklmn}$$

where M is a component common to all sources of variation ;

A_i is a component common to all pupils in school type i ;

B_j is a component common to all pupils in school size j ;

C_k is a component common to all pupils of sex k ;

$(AB)_{ij}$ is a component resulting from the interaction between school type i and size j ;

$(AC)_{ik}$ is a component resulting from the interaction between school type i and sex k ;

$(BC)_{jk}$ is a component resulting from the interaction between school size j and sex k ;

$(ABC)_{ijk}$ is a component resulting from the interaction between school type i, size j and sex k ;

D_{ijkl} is a component common to all pupils in the l^{th} school within school type i, size j and sex k ;

E_{ijkm} is a component common to all pupils in the m^{th} choice within school type i, size j and six k ;

$(DE)_{ikjlm}$ is a component resulting from the interaction of the l^{th} school and m^{th} choice within school type i, size j and sex k ;

and e_{ijklmn} is a component specific to the nth pupil in the l^{th} school and the m^{th} choice within school type i, size j and sex k.

i, j, k and m take on the values 1 and 2 ; l takes on the values 1, 2, . . . 4 ; and n takes on the values of 1, 2, . . . 6. It should be noted that school type,, size and sex are all crossed with each other, and that schools and choices are also crossed with each other, both being nested within school type by size by sex. The breakdown of the total variance, together with the degrees of freedom and mean square expectation of each effect, is as shown below. It is assumed that there are a, b and c categories of A, B and C ; that there are d and e categories of D and E within each of the abc categories of A, B and C ; and that there are n pupils within each of the abcde categories of D and E. (Thus, in the present study, a, b, c and e all equal 2, d=4 and n=6). It must be noted too that A, B, C and E are fixed effects, while D is a random effect.

Source of Variation	Degree of Freedom	Mean Square Expectation
School Types (A)	a −1	$\sigma^2+en\sigma_D{}^2+bcd\sigma_A{}^2$
School Size (B)	b −1	$\sigma^2+en\sigma_D{}^2+acd\sigma_B{}^2$
Sex (C)	c −1	$\sigma^2+en\sigma_D{}^2+abd\sigma_C{}^2$
AB	(a −1) (b −1)	$\sigma^2+en\sigma_D{}^2+cd\sigma_{AB}{}^2$
AC	(a −1) (c −1)	$\sigma^2+en\sigma_D{}^2+bd\sigma_{AC}{}^2$
BC	(b −1) (c −1)	$\sigma^2+en\sigma_D{}^2+ad\sigma_{BC}{}^2$
ABC	(a −1) (b −1) (c −1)	$\sigma^2+en\sigma_D{}^2+d\sigma_{ABC}{}^2$
Schools (D), within ABC	abc (d −1)	$\sigma^2+en\sigma_D{}^2$
Choices (E), within ABC . .	abc (e −1)	$\sigma^2+n\sigma_{DE}{}^2+dn\sigma_E{}^2$
DE, within ABC	abc (d −1) (e −1)	$\sigma^2+n\sigma_{DE}{}^2$
Pupils, within DE	abcde (n −1)	σ^2
Total	abcden −1	

(*Manuscript received* 16*th March*, 1967)

A COMPARISON OF ESSAY GRADES OBTAINED BY SCIENCE AND ARTS GRADUATES IN A DEPARTMENT OF EDUCATION

By P. E. RICHMOND

(*Department of Education, University of Southampton*)

SUMMARY. The essay grades obtained by four-hundred honours graduates in a university department of education were examined to investigate differences in essay-writing performance between arts and science students. Significant differences exist between mean grades of the different groups but also between their classes of degree. An analysis of covariance to hold degree class constant was carried out and the differences between the essay grades of science and arts students became insignificant.

I.—INTRODUCTION.

KIRKMAN (1966) tells us that " it is the exception rather than the rule to find a young engineer or scientist who can write plainly and readably, on specialist or general topics," and the *New Scientist* (1966) reports that " Professors have lately made sad complaint that 90 per cent. of our young engineers and scientists are incapable of expressing themselves clearly and concisely on paper." Subjective judgments are frequently made about the poor quality of the writing of scientists and the comments are often directed at undergraduates. Firm evidence that their writing is markedly inferior to that of their fellows studying non-scientific subjects is, however, difficult to come by.

A university department of education offers a unique opportunity to compare abilities of graduates in different disciplines. It recruits potential teachers qualified in a wide range of subjects and submits them to a one-year course (fully described by Wagner and Douch (1965)) in which only a small minority of time is spent in studies directly related to their degree subject. The rest is new and common to all. The aim of the course is to examine, criticise and develop ideas, " to have a liberalising rather than restrictive effect on the student's outlook," and some education departments, of which Southampton is one, assess their students by means of essays written throughout the course rather than by conventional written papers. The whole procedure would seem to be biased against the mathematics or science graduate who is so often thought of as unable to cope with ideas and writings outside his own narrow field. It does, however, offer a measure of the ability of the scientist to compete with the arts man in writing continuous prose.

Five essays (see Wagner and Douch) are assessed : (1) Method—an account of the theories and practices of teaching a particular subject ; (2) History and Practice of English Education; (3) Principles of Education; (4) An essay on any subject of educational interest, chosen by the student ; (5) A critical account of the student's teaching practice, which occupies the whole of the middle term of a three-term course. In every case the student can choose from a wide range of topics or treatments and the marking favours imaginative, critical discussion over any recital of facts or second-hand opinion. The markers, in the first instance, are department staff and, except for the method essay (where a subject specialist marks the essays written by graduates in his own subject), the essays are assessed by members of academic staff drawn from both the arts and the science sides. The external assessors, not chosen for their degree subject, check the quality of the essays before a final grade is arrived at. All of the intercorrelations among the five essays' grades are significant at the 1 per cent. level.

II.—METHOD.

To investigate the differences in performances of graduates in the sciences and the arts, essays written between October, 1960, and July, 1965, were examined. They had been graded A to E, qualified by + or − and for statistical purposes were allotted a score from 15 (A+) to 1 (E−) and the scores for all an individual's essays were totalled. Consideration was given only to honours graduates in a single subject to avoid the blurring of subject group boundaries which multiple-subject degrees would have caused and the essays were divided into subject groups on a science-arts continuum as listed in Table 1.

Table 1 shows that there *is* an increase in average essay grade from the mathematicians to the foreign language and English graduates, but it also shows clearly the superiority of the degree class of the arts people. An analysis of variance showed that differences in essay grades were significant between subject groups. The better essay writers were in the groups with the better degrees. The differences between groups might well be due to more able arts than science graduates choosing to teach. In consequence, it was decided to carry out an analysis of covariance so that essay grades could be compared, with degree class held constant. This analysis (Table 2) showed that the differences between the means of the subject groups were no longer significant and Table 3 shows how small these differences become. Provided we allow for the superior degree class of the arts graduates, we must reject the hypothesis that differences in essay-writing ability exist between arts and science students in a department of education. Caution is required, however, in the case of the mathematics graduates whose grades, even after correction, are substantially below the others.

III.—DISCUSSION.

The general conclusion cannot be applied to all graduates for at least two reasons. The sample of graduates choosing to teach may not be representative of graduates as a whole and observation and discussion lead us to suspect that the scientists do not take essays in their stride but work hard at mastering a relatively unfamiliar task.

Young (1965), Bruner (1966) and Hudson (1966) are only three of many writers who have recently stressed the inadequacy of tests given out of context or in a situation in which the subject has little interest. Kirkman's (1967) vocabulary tests have shown substantial differences among different disciplines but, as the comments of some of his subjects show, the attitudes the undergraduates bring to the tests are as varied as their scores.

In the experiment reported here, evaluation and education proceeded simultaneously and scientists tackled their essays as part of the course with an interest matching that of the non-scientist students. Their performances were similar and, in view of their similar motivation and involvement, this is less surprising than might at first appear.

TABLE 1

ESSAY GRADES AND DEGREE CLASS AMONG GRADUATES IN A DEPARTMENT OF EDUCATION 1961—1965, INCLUSIVE).

	Essay grades			Degree class	
	N	M	S.D.	M	S.D.
Mathematics	49	46·45	6·11	3·08	0·75
Physical sciences	55	48·04	6·42	3·16	0·87
Biological sciences	31	48·26	6·68	3·03	0·74
History, geography, economics	102	48·83	5·84	2·93	0·73
Languages	104	50·27	6·85	2·72	0·73
English	59	49·17	7·04	2·81	0·70

Essay grades : Sum of five grades (A+ =15, 1=14, . . . E=2, E− =1).
Degree class : 1st class honours=1, Upper 2nd class honours=2, Lower 2nd class honours=3, 3rd or pass=4.
Analysis of variance :
Essay grades F ratio=2·56 $p<0{\cdot}05$ (D.f. 5,394).
Degree class F ratio=3·45 $p<0{\cdot}01$ (D.f. 5,394).

TABLE 2

ANALYSIS OF COVARIANCE : THE VARIATION OF ESSAY GRADES BETWEEN SUBJECT GROUPS WITH DEGREE CLASS HELD CONSTANT.

Adjusted essay scores.

	Between	Within	Regression	Total
Adjusted sum of squares	297·2	15,567·2	1,437·2	17,301·6
Degrees of freedom	5	393	1	399
Variance estimates	59·4	39·6		

F ratio=1·51 $p>0{\cdot}05$

TABLE 3

MEAN ESSAY GRADES ADJUSTED FOR DEGREE CLASS.

Group	Adjusted mean
Mathematics	46·8
Physical science	48·6
Biological science	48·5
History, geography	48·9
Languages	49·8
English	48·9

IV.—REFERENCES.

BRUNER, J. S. (1966). *Toward a Theory of Instruction.* Harvard University Press.

HUDSON, L. (1966). *Contrary Imaginations.* London : Methuen.

KIRKMAN, A. J. (1966). The deficiencies in students (*a*)linguistic, in *Proceedings of the Conference on the Training of Scientists and Technologists in Written Communication.* R. G. Martyr, Kodak, Ltd., Ruislip, Middlesex.

KIRKMAN, A. J. (1967). Command of vocabulary among university entrants. *Educ. Res.*, 9, 151-159.

New Scientist (1966), 30, 504.

WAGNER, F. W., and DOUCH, R. (1965). The education of graduate teachers. *Education for Teaching*, 65, 24-36.

YOUNG, M. (1965). *Innovation and Research in Education.* London : Routledge and Kegan Paul.

(*Manuscript received, 3rd May, 1967*)

THE EFFECT OF URBANIZATION UPON SCIENTIFIC CONCEPT ATTAINMENT AMONG HAUSA CHILDREN OF NORTHERN NIGERIA

By H. E. POOLE

(*Dalhousie University, Halifax, Nova Scotia*)

SUMMARY. An attempt is made in this paper to measure the influence of varying types of settlement upon the attainment of scientific concepts among African school children. An adaptation of King's (1963) Piaget-type schedules was administered to 150 school children of the Hausa tribe of Northern Nigeria in eight groups drawn from very isolated rural localities, from city centres, and from areas of intermediate status. The results strongly confirmed the expected superiority of the urban children. Finally, an English version of the test was used upon a group of English school children to provide a cross-cultural comparison. The significance of the results for science education is briefly discussed.

I.—INTRODUCTION.

PIAGET and his imitators have produced a connected and plausible account of intellectual growth from birth to adolescence. Piaget himself in the early days seemed to assume that mental progression was primarily a reflection of maturation of the central nervous system. Since then, however, he and others have become increasingly aware of the part played by properly timed experience in the emergence of successive intellectual structures. But different cultures provide their members with different experiences, inevitably, because of the widely varying conceptual systems they possess. It is hardly surprising, therefore, that the rate and character of intellectual growth should show considerable cross-cultural variation when measured against the yardstick of the Piagetian stages, which imply mastery of concepts central to Western culture. Nevertheless, it is difficult to see how the less advanced societies can achieve the high living standards at which they aim without assimilating large portions of the Western conceptual system, not least those concepts of scientific significance with which this study deals. The Piagetian model performs the necessary, if sometimes depressing, task of indicating basic intellectual hindrances to national development. It was largely in accordance with the spirit and procedures of this model that the present study was performed.

II.—PREVIOUS RESEARCH.

Nearly all cross-cultural studies of concept attainment have suffered from one of two elementary though serious defects. The first is the outcome of assuming that because one test has been modelled on another, or inspired by it, the results of each may be validly compared. Plainly, this will not do. The criteria for validity in this context demand more than can be satisfied by subjective assessments on the apparent similarity of tests. If legitimate cross-cultural comparisons are to be made, the tests and other experimental conditions must correspond fairly exactly across the samples of the cultures to be compared. Stripped of jargon, this is common sense; but researchers using Piagetian methods have seemed especially prone to forget it. And, clearly, there is in many cases a good reason for this refusal to produce evidence other than that of a single exotic sample, namely, that in adapting a mode of presentation to the peculiarities of a cultural situation, an experimenter may well find himself with a test that is unusable elsewhere. Such tests can promote insight; they cannot provide comparisons. It is to avoid this criticism that the present study uses only items on which the performance of Western children is well established; even so, the responses of a group of English children to this particular presentation and combination of the material are given.

The second defect originates in a failure to recognize that variations between cultural sub-groups in non-Western cultures may be far wider than occur in Western society. If this possibility is ignored there is a danger of basing cross-cultural comparisons on untypical samples. The urban and rural samples reported upon in this study embody wide intra-cultural variations; they are immeasurably further apart in cultural terms than town and country in Western Europe. The present study is concerned with measuring changes occurring within an African society, rather than with comparisons between African and European children. Specifically, it aims to show how scientific concept attainment in Hausa society is affected by degree of urbanisation.

III.—SAMPLES AND TESTS.

To make the comparison in concept attainment between urban and rural children in Northern Nigeria, samples were drawn from eight schools, three of which were located in a city, three in isolated rural areas, and two from large main road market villages intermediate between these extremes. The rural and intermediate samples were taken from a radius of thirty miles from the city. All the subjects were 10–11-year-old children of the Hausa tribe of Zaria Province, Northern Nigeria. Every attempt was made to ensure that the children tested were in fact within the specified age-range; but since records of birth have not hitherto been kept in these parts, it is probable that a few children from adjoining age groups have been included.

The peculiar circumstances which must be reckoned with in such new countries as Northern Nigeria made it necessary to vary the size of the samples. Four of the samples were of fifteen children, three were of twenty and one of thirty; but each sample was randomly selected from the target age-group within each school. The only departure from the random principle was in the exclusion of six defectives (from the whole original sample of 150) who were unable to make the very simple responses required by the schedule. As no social or other bias affects primary school admissions, the school children themselves can be presumed to represent all levels of Hausa society.

Finally, a group of forty English children (I.Qs. 90–110) was tested on a straight translation of the Hausa test, employing the same illustrations and equipment. The sample was randomly drawn from a very large junior school in the suburbs of a Southern English city. All the children came from estate-type homes of lower professional or working class status. This English sample has been included on the assumption that it would have a degree of familiarity with scientific concepts towards which the Hausa group is moving.

As a measure of scientific concept attainment a shortened version of King's Schedules (1963) was used—shortened because not all items in King's original could be successfully rendered in Hausa. The thirty-four items of the test fell into the following categories:

Category	Item	Total Items
1. Measures (time and space)	1, 2, 3, 4, 6, 7.	6
2. Prediction of spatial relationships..........	5, 8, 11, 12, 15, 18, 19	7
3. Conservation of solids and fluids	9, 10, 13, 14, 16, 17	6
4. Grasp of mechanical principles	20—25............	6
5. Attribution of life (including Animistic tendencies)	26—34............	9

A summary of the test items is to be found in the Appendix.

There is considerable overlap in content between categories 2 and 3; numbers 11 and 12, for example, might be regarded as involving conservation concepts. They do; but in common with all the items in category 2, they presuppose ability to interpret diagrams and predict appearances in two dimensions. It would be difficult to maintain the distinction between these two categories in the face of a determined critic; nevertheless, in view of the special difficulties which Africans find in the perception of space—as noted by Hudson (1960), Cryns (1962), and others—the distinction has been made, though tentatively.

Great care was taken to ensure that the test was fully comprehensible and the final version was arrived at only after prolonged experiment and revision. All groups were given a supervised standard practice in using the answer sheet. Practice material was unconnected with the content of the subsequent test. Responses were limited to 'yes,' 'no,' or a capital letter written in the answer space. Instructions and questions were played from sound recordings supported by illustrations. Any illegible answers were checked with the child immediately after the test. (This proved necessary only for seventeen of all answers throughout the series of tests.) The Hausa samples were tested late in 1965, the English sample early in 1966.

IV.—RESULTS.

(*a*) *General.*

Despite some prior misgivings about the possible behaviour of the rural subjects under test conditions, the children in all samples responded well. The frequency of correct responses and the rank order of item difficulty within each sample are set out in Table 1.

Most of the results were in the expected direction. Comment will here be limited to apparent anomalies. The higher Hausa scores on item 19 probably reflect life in a tropical environment, where location of shade is necessary to comfort, even survival; as for item 21, daily chores of carrying are familiar—and onerous—to Hausa children. It was found that the results on item 7, which reverse the usual Hausa rural-urban scoring order, reflect special knowledge acquired by one of the rural samples just before the test. Most surprising, surely, was the absence of any significant cross-cultural variation in items 26–34 which were intended to ascertain tendencies towards animistic thinking. However injuring to European vanity this persistence of animism may be, it is now fairly well established; Dennis (1953) found traces of it even among college and university students.

(*b*) *Category and Item Difficulty.*

A further look at the rank order of difficulty presented in Table 1 shows an unexpected concordance between the samples in rank order difficulty of the categories (Table 2).

This correspondence extends to the individual items within the categories, as Table 3 shows. With the exception of the low correlation in the spatial category, which was expected, the Hausa order of individual item difficulty seems fairly similar to that of the English group.

TABLE 1

Percentage Correct Responses (PCR) and Rank Order Item Difficulty (RID).

	Nigerian Subjects						English Subjects	
	Rural (N 50)		Intermediate (N 30)		Urban (N 70)		(N 40)	
Item	PCR	RID	PCR	RID	PCR	RID	PCR	RID
1	10	4	17	6	16	4	57	10
2	66	30	73	29·5	74	27·5	95	28·5
3	28	18	43	15	46	13·5	80	22
4	8	3	23	9	26	6	38	3
5	12	5·5	13	3·5	23	5	42	6
6	2	1	3	1·5	10	1·5	40	4·5
7	14	7·5	17	6	11	3	17	1
8	12	5·5	3	1·5	10	1·5	40	4·5
9	48	26·5	60	24	73	24·5	75	18·5
10	24	14	40	12	39	11	47	7
11	30	19	43	15	46	13·5	92	27
12	18	10	30	10	41	12	97	31·5
13	26	16·5	43	15	59	18	72	17
14	20	11	20	8	31	8	60	11
15	22	12	53	21·5	34	10	95	28·5
16	24	14	53	21·5	57	17	77	20
17	46	24·5	63	25	74	27·5	90	25·5
18	54	28	53	41·5	69	22	97	31·5
19	46	24·5	50	19	63	19	35	2
20	16	9	13	3·5	30	7	62	12·5
21	74	31	77	31·5	90	31	80	22
22	24	14	43	15	53	15	70	16
23	14	7·5	17	6	54	16	62	12·5
24	56	29	43	15	74	27·5	90	25·5
25	36	22	37	11	70	23	80	22
26	6	2	47	18	33	9	50	8·5
27	82	32	67	26·5	91	32	97	31·5
28	88	33	100	34	100	34	100	34
29	26	16·5	53	21·5	67	20·5	65	14
30	48	26·5	77	31·5	77	30	75	18·5
31	44	23	73	29·5	67	20·5	50	8·5
32	32	20·5	70	28	73	24·5	67	15
33	32	20·5	67	26·5	74	27·5	82	24
34	92	34	97	33	96	33	97	31·5

TABLE 2

Rank Order of Category Difficulty (By Sample).

	Rural	Inter	Urban	English
Measure	1	1	1	1
Spatial	2	2	2	3
Conservation	3	4	3	2
Mechanical	4	3	4	4
Animism	5	5	5	5

TABLE 3

Correlation in Rank Order of Item Difficulty, By Category.

Correlation between—	Rural-Inter	Inter-Urban	Urban-English
Measure	·780	·985**	·715
Spatial	·839*	·722*	·340
Conservation	·838*	·943**	·772
Mechanical	·786	·757	·814
Animism	·844**	·758*	·954**

* Significant at ·05 level. ** Significant at ·01 level.

(c) *Frequency of Correct Responses.*

Analysis of variance shows the differences between scores within the Hausa group and between Hausa and English samples, to be significant in the predicted direction beyond the ·01 level. It is less easy to achieve reliable significances when analysing data at the category level, at which the power of statistical tools is low, and interpretation of the data must lack the degree of certainty one would wish for. When the Whitney-Mann U-test was applied to the ranked scores in two categories at a time, only three differences emerged that were significant at the ·05 level. Between the Rural and the Intermediate groups these were with the Conservation and the Animism items; with the Urban-English comparison, the significant difference was with the Spatial items. Finally, when the English scores were expressed as quotients (X 100) of the composite average Hausa scores, the results were: Measure (205), Spatial (202), Conservation (158), Mechanical (151), Animism (115). These English quotients are significantly higher at the ·05 level in all but the Animistic category.

V.—Discussion and conclusion.

It has been assumed that the significant differences noted between Hausa samples and between Hausa and English groups are the reflection of cultural differences. These differences appear to be a measure of the extent to which the concepts of Western society—as presumably represented in the English group—have been superimposed upon the traditional culture of the Hausa group.

The results contained in the previous section suggest ways in which this cultural change might be construed. In the first place, it is remarkable that there should be so high a degree of concordance in the rank order of difficulty (Tables 2 and 3). The presence of common difficulties certainly argues the penetration of Western ideas into Hausa society. This interpretation is reinforced by the divergence of Hausa and English scores with increasing category difficulty, noted in section (c) above. Such a divergence, which distinguishes between rank and actual difficulty, would be expected in an acculturative situation. Clearly the more expert will have greater facility on the more difficult items than the less expert.

It can be argued, of course, that the order of category difficulty is not only common, but culture-free. There is no *a priori* reason to doubt that some operations may be intrinsically more difficult than others in terms of inherent human capabilities; in fact, the sequential appearance of the Piagetian stages presupposes an order of difficulty that is independent of previous learning.

However, the fact that an order of difficulty can be culture-free must not be taken to mean that the Piagetian stages are a culture-free test of general intelligence. The part played by general intelligence in cross-cultural variations of the stages cannot be assessed. All that can be stated with confidence at present is that the categories measure abilities which societies encourage to varying degrees or not at all.

The conceptual dominances of the Hausa are changing and the results of this study confirm the acculturative role of urbanisation. Throughout history cities have been focal points of change, and in the twentieth century it has been from the cities that the white man's civilisation has radiated. The paraphernalia of imported culture—schools, factories, airports, clinics, clocks, and so forth—tend to be concentrated in urban areas. Town dwellers, therefore, are more likely than villagers to encounter and use scientific concepts.

The conservatism of rural communities on the other hand ensures slower acculturation. As a rule, these communities are intolerant of innovation either from within or from without—and scientific ideas are new ideas. It seems that the cohesiveness which enables rural groups to preserve their outward customs also promotes the survival of their conceptual systems. Conversely, urbanisation forces upon people some degree of receptiveness to novelty. But the fatal flaw in the old Hausa culture was its inherent inability to adapt. It had been slowly evolved to meet a complex of fairly rigidly defined relationships which are now only to be found in villages. In the city the ways and assumptions of the old society have been put to the test and found wanting. The bewildering situation in which the urban Hausa find themselves, and the need to master the problems generated by Westernisation, account for their eagerness to learn new social and intellectual skills.

Yet the traditional culture, though losing ground, hangs on to life in the many institutions that have survived the impact of urbanisation. The fact that most Hausa women are illiterate, for example, and grow up to regard their ignorance as a womanly virtue, guarantees the transmission of ideas and attitudes which must impair the ability of children to assimilate Western concepts.

The pre-school and out-of-school education of Hausa children, then, provides little in the way of scientific concepts, which are the currency of scientific thinking and learning. Moreover, the capacity of formal education to make good this deficiency is limited. For one thing, school learning represents more or less of a discontinuity in the intellectual life of Hausa children; for another, studies on both animals and human beings suggest that successful assimilation of many concepts depends upon appropriate stimulation during the earlier years of life. These factors must, in part at least, account for the difficulties experienced by Hausas in learning science.

VI.—REFERENCES.

CRYNS, A. G. T. (1962). African intelligence: a critical survey of cross-cultural intelligence research in Africa South of the Sahara. *J. Soc. Psychol.*, 57, 283-301.

DENNIS, W. (1953). Animistic thinking among college and university students. *Scient. Monog.*, 76, 4, 247-9.

HUDSON, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. *J. Soc. Psychol.*, 52, 183-208.

KING, W. H. (1963). Attainment of scientific concepts among children. *Brit. J. Educ. Psychol.*, 33, 240-252.

(*Manuscript received* 30th *March*, 1967)

VII.—Appendix.

Summary of Scientific Concept Attainment Test Adapted from King (1963).

1.—Estimate time between two taps (15 seconds ; accept 10–20).

2.—Day of week.

3.—Day of week in . . . (nearby town).

4.—Estimate length (12–18 feet ; accept plus or minus one-sixth).

5.—Which trajectory gives ball thrower greatest distance ?

6.—Time between taps (45 seconds : accept 35-55).

7.—Estimate distance (about 50 yards ; accept plus or minus one-sixth).

8.—Which trajectory of jump gives greatest distance ?

9.—Stone added to water. Is level higher, lower, or same ?

10.—Sand added to water. Is level higher, lower, or same ?

11.—Water poured into jar of different shape. Predict new level of water.

12.—As above, but with jar of different shape.

13.—Water in tall narrow jar poured into squat, broad basin. Is there more water or less or the same ?

14.—As for 13, but is water heavier or the same weight ?

15.—Predict level of water in tilted jar.

16.—Does change in shape of plasticine make it heavier ?

17.—As above, but is there more material as result of change ?

18.—What will house built on the side of a steep hill look like ?

19.—Man and lamp. Predict direction in which shadow will lie.

20.—See-saw. Where must unequal weight be placed to balance ?

21.—Two men carrying long pole. Who carries the heavier weight ?

22.—Gears. Three interlocking wheels of differing size. In which direction does last one turn ?

23.—As above. Which turns fastest ?

24.—Wheels of unequal size with belt, passing round both. In which direction does second wheel turn ?

25.—As above, but in this case, the belt is crossed between the wheels.

26.—Is sun alive ?
27.—A tree ?
28.—A dog ?
29.—A fire ?
30.—A candle ?
31.—A river ?
32.—A train ?
33.—An aeroplane ?
34.—A bird ?

THE EFFECT OF WESTERN EDUCATION ON SCIENCE CONCEPTUALISATION IN NEW GUINEA

By J. R. PRINCE, M.A.

(*Warden of the Preliminary Year, University of Papua and New Guinea*)

Summary. A pilot survey of the development of basic physical science concepts was undertaken in three areas of the Territory of Papua and New Guinea. In the course of the survey certain special groups were tested and these results, together with those of the main survey, have been examined for the light they throw on the effect of environmental and inheritance factors on conceptual development.

The three diverse areas of the Territory yield similar results when test performance is analysed by school grade but there are differences between the areas when the results are analysed by calendar age. It appears that school grade is more significant than calendar age in determining conceptual development in the Territory, and this suggests that educational environment may be a major factor in conceptual development in the Territory. This conclusion is supported by the performance of a teachers' college group which had had special instruction regarding substance and weight and a primary school which had undertaken a comprehensive programme of experience in measurement.

The performance of a group of students at a non-English medium teachers' college reveals significantly poorer development in all test items requiring the concept of conservation. Statistical analysis of these test results against years of total schooling and years of English education reveals that English education is more significant in relation to physical concepts than total years of schooling.

The four lines of evidence considered in the paper all point to the importance of environmental factors, particularly those connected with certain aspects of education, in developing concepts which involve the conservation of physical quantities.

I.—Introduction.

Studies on conceptualisation in cultures other than that of the industrialised Western countries have raised a number of important issues. Two of such issues were highlighted at the 1967 Congress of the Australian and New Zealand Association for the Advancement of Science. The fact that the Piagetian stages of concept development are found in non-Western cultures, even though conservation of physical quantities is achieved some years later than in Western cultures, raises the question as to how far such development is the product of intruding Western influence, and how far it would occur in the undisturbed indigenous culture by the simple maturational process of development. The second and related issue is that which concerns the degree to which inherited factors determine conceptual development.

Price Williams (1962), in his study of classification among children of the Tiv tribe in Nigeria, indicates that there is little difference between the performance of African and European children, provided the test material is of a nature familiar to the culture in which the test is being carried out. A similar inference could be drawn from the study of Goodnow (1962) in Hong Kong. Greenfield (1966), in a study of conservation among Wolof children in Senegal, finds that "In terms of grade level, the Senegalese figures are close to being identical to the Western ones, although the Wolof children are behind in terms of chronological age. The parallel findings certainly cast strong doubts on any simple maturational notion of development." This is supported by the writer's findings in New Guinea, set out in this paper.

These findings are not claimed to arise out of an investigation specifically into the points now under discussion. However, any cross-cultural study of conceptualisation must bear upon them and certain results from a conceptualisation survey undertaken in three different areas of the Territory in 1966 do throw some light on these important issues.

II.—THE SURVEY.

The survey referred to involved 2,700 school children from Primary Grade 3 to Form III of High School in three distinct geographical and cultural regions of the territory of Papua and New Guinea. The survey was conducted by administering a group test of twenty-three items, and by follow-up interviewing where possible to check on the intelligibility of the questions, reasons for unexpected responses, and thought patterns behind both correct and incorrect responses. (Examples of test items are given in the Appendix: the full list may be had from the author.) The test items were demonstrated to classes and the children were required only to write down the capital letter corresponding to the answer of their choice, selected usually from three answers written on the blackboard by an assistant. The test items covered concepts of length and displacement, conservation of quantity, weight, area and volume and the view of density in relation to the nature of the substance.

The responses of the children together with coded information about each child, were processed on a computer to show the percentage of correct responses in each class grade, and in each calendar age group, both in the whole sample and in the separate districts tested. Two schools were singled out for individual processing and one or two special samples, particularly relevant to this paper, were also processed.

The responses both in the test and in the interviews show the expected pattern of Piagetian stages, though conservation is not achieved until much later than in Western European culture. An estimate of this difference in terms of calendar age is made difficult both by the uncertainty of age statements in the Territory, and also, perhaps more significantly, by the considerable variations between the districts tested in the sample. Yet, when performance is measured against school grading, there is no great difference among the three districts tested (Figures 1-4).

The three districts selected differ very considerably in almost every respect except the educational programme now being provided. The Central District contains Port Moresby and is a coastal area of long European contact. The schools tested in this area had all been in existence for a considerable period, and were attended by virtually all children of school age in the surrounding area. Hence, school grades were relatively homogenous for age with Standard 3 averaging about 9 years of age, Standard 4 10 years and so on. The Highlands district on the other hand,, is an area of recent contact where education is a recent innovation. Several of the schools in the sample had commenced so recently that they had not yet taken a group to Standard 6. Ages were therefore extremely variable, with Standard 3 including age ranges from 9 to 13 and higher grades similarly. The Western District sample also showed a considerable age spread within school grades.

Fig. 1 CONSERVATION OF SUBSTANCE BY SCHOOL GRADE

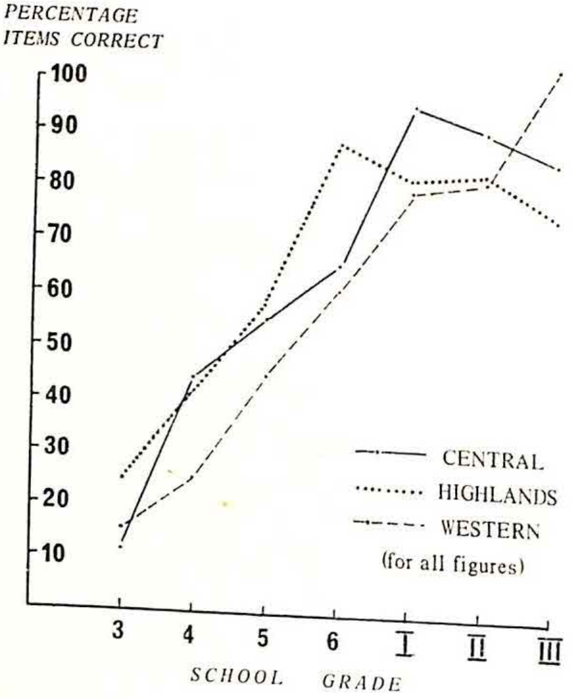


Fig. 2 CONSERVATION OF SUBSTANCE BY CALENDAR AGE

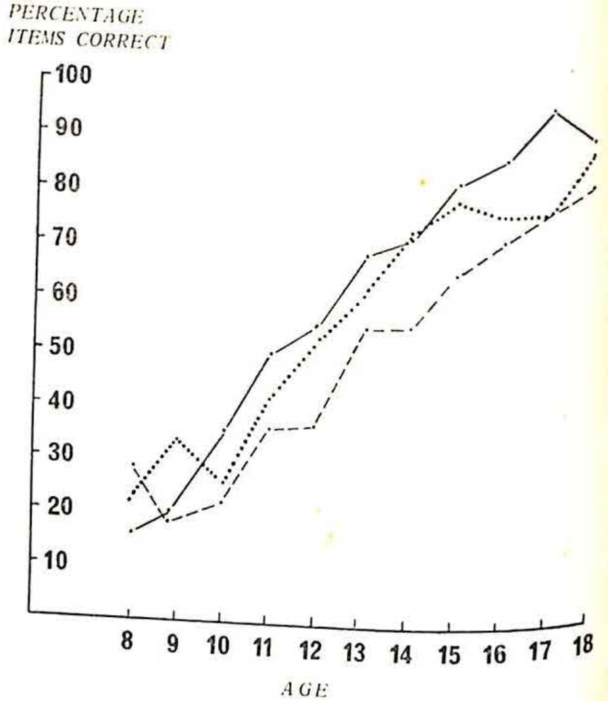


Fig. 3 CONSERVATION OF AREA BY SCHOOL GRADE

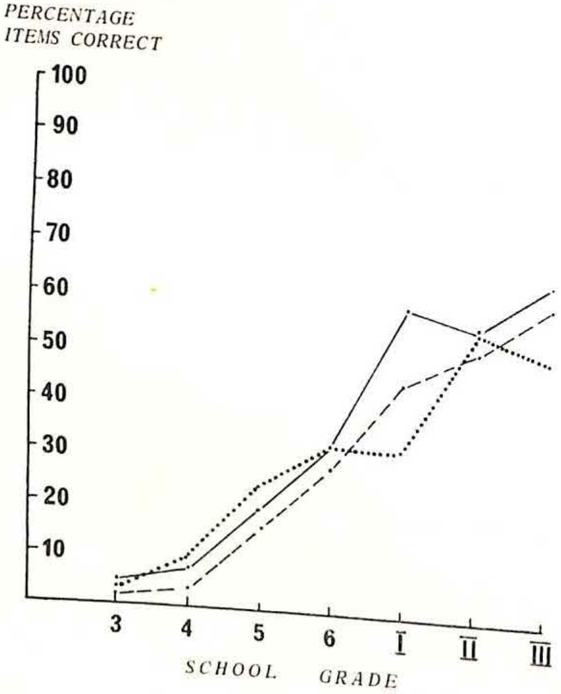
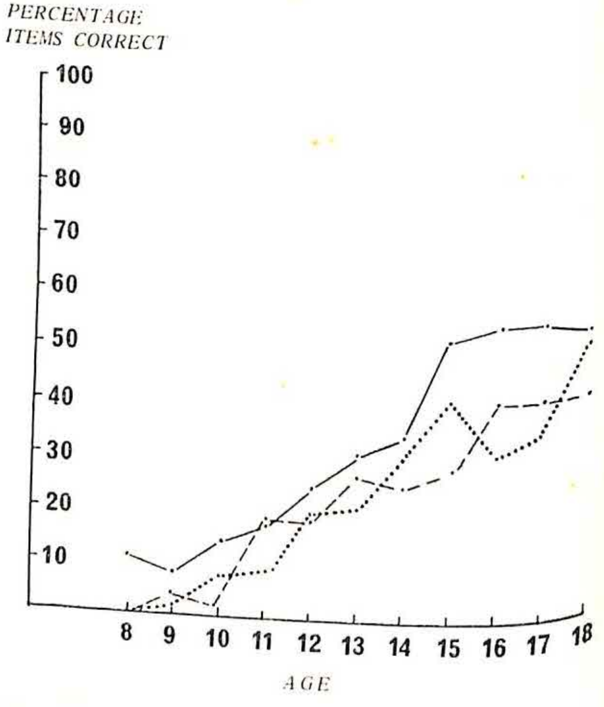


Fig. 4 CONSERVATION OF AREA BY CALENDAR AGE



Comparison of the grade performance and age performance curves reveals two features:

(*a*) As assessed by grades the performance of the Highlands is comparable with that of the Central District, while the Western Districts' primary schools perform a little less well than either. This last result is explicable on grounds of motivation in an area where opportunities for advancement are very limited. The age curve of the Highlands is at almost all points below that of the Central District. It could be interpreted that the Highlands child is less intelligent than the central coastal Papuan, though this seems unlikely to the author. It seems more probable that the poorer Highland performance on the age scale is due to the fact that for a given age sample the average school grade in the Highlands is lower than on the central coast and that performance in this test is more directly related to school grade than to calendar age.

(*b*) A comparison of the districts' age-performance graphs shows that the Central Districts' performance rises fairly steadily across the graph, while the other two graphs show more irregularities. Bearing in mind that the Central District sample showed a high correlation between grade and age, whereas the other samples did so to a lesser extent, it is again clear that performance on the test was more closely related to school grade than to calendar age.

Insofar as the test was a valid measure of conceptual development, it would therefore appear that Western education influence is the important factor in developing the concepts involved in the test. This opinion is confirmed by the results of a number of special groups tested in the course of the survey.

III.—SPECIAL GROUPS TESTED.

Four special groups tested in the course of the survey threw some further light on the question of how far cultural factors influence concept development. The four groups were:

(*a*) *Siokiei Primary 'T' School.* This is a primary school, working on the territory syllabus (hence 'T') which had been used to develop an experimental programme of modern mathematics based largely on the work of Dr. Z. P. Dienes. Most grades in the school had had several years experience of work in set theory, classification and discriminatory work with attribute blocks and Venn diagrams, and other exercises in mathematical logic. The school had, however, done little work in measurement as the school was tested before the measurement section first introduced in 1966 had been undertaken. There was no Standard 6 in the school.

(*b*) *Mt. Hagen Primary 'T' School.* This school was also working on a modern mathematics programme although the programme had been working for a shorter time than at Siokiei, and had developed along slightly different lines. In particular, the programme used had contained an extensive section on simple measurement which included estimating games, such as how many of tin A would fill tin B, and so on. This school was tested at the end of the year, after the measurement programme had been completed in all grades. Unfortunately, in order to test at the end of the year, the testing had to be delayed until the completion of the final primary examination and all but six of the Grade 6 children had returned home to their villages. This fact, and the absence of a Grade 6 at Siokiei explains why Grade 5 performances had been used for comparison in Table 2.

(*c*) *Awaba Teachers' College.* This college, conducted by the Unevangelised Field Mission, located in a fairly remote area of the Western district of Papua, consisted of forty-four teacher trainees made up as follows :

Undertaking one-year A course after Standard 6 ..	10
Undertaking one-year A course after Form I	29
Undertaking first-year of B course after Form II ..	5
	44

(*d*) *Rintebe Teachers' College.* Conducted by the Lutheran Mission, near Goroka in the Eastern Highlands district, the college contained 111 teacher trainees being taught in Pidgin and Kate (a vernacular from the Finschhafen district) to teach in those media. The test in this case was administered in Pidgin which was well understood by the Kate section of the College also. There was a wide variation in educational background among the 111 trainees, some having had no formal education in English and others a number of years (Table 1).

TABLE 1

EDUCATIONAL BACKGROUND OF STUDENTS AT RINTEBE T.C.

Formal Education : Years	0	1	2	3	4	5	6	7	8	9	10	11+	
No. of students	5	2	2	2	9	7	14	11	22	10	15	12	Total 111
English Education : Standard..........	0	1	2	3	4	5	6						
No. of students	42	6	11	13	13	9	17						Total 111

IV.—COMPARATIVE TEST RESULTS.

Section A of Table 2 compares the results of Rintebe Teachers' College and Awaba Teachers' College, with the whole Form I sample tested in the survey. This comparison has been made because most of the students at Awaba had completed Form I but had not undertaken any further high school education with its associated work in higher mathematics and science. They had come mostly from the lower half of the previous year's Form I, the better students having proceeded to Form II.

The generally good performance of Awaba on weight and quantity concepts probably reflected the fact that the staff of the College had recently given the trainees an extensive programme of work on weight estimation and ordering. The programme threw similar emphasis on length estimation and ordering. It is not clear why a corresponding improvement was not achieved in the length-displacement group.

TABLE 2

SCIENCE CONCEPT QUESTIONNAIRE: PERCENTAGE CORRECT RESPONSES.

Topic	No. of items	A—Teacher Training Group			B—Grade Five Group		
		Rintebe T.C.	Awaba T.C.	All Form I	Mt. Hagen Gr. 5	Soikiei Gr. 5	All Gr. 5
Length—Displacement	4	7	16	33	35	19	16
Conservation of Weight	2	45	82	69	80	35	56
Conservation of Quantity (plasticine)	3	22	71	56	70	29	46
Conservation of Quantity (water)	2	40	80	85	82	57	56
Conservation of Area	4	4	41	44	29	19	20
Conservation of Volume	2	19	55	64	82	35	44
Density Ideas	2	37	43	53	47	12	39
Stone in Water	2	93	100	98	96	92	92
Wood in Water	2	34	28	30	15	11	30
Number in Sample		105	44	423	34	31	518

On all test items involving conservation of physical quantities, the performance of the students at Rintebe is so much below that of the other two groups that one must conclude that conservation of physical quantities is not developed by even a long period of education in Pidgin or Kate. In two groups of items, the Rintebe students score comparably with the others. These enquire about the effect of immersing a stone in water and of placing a piece of wood in water, the first question asking whether the object will sink and the second what it will do to the level of the water. These two groups involve conservation only to the extent of realising that a space cannot be occupied by two objects at once. It would seem that experience in the indigenous culture does bring about this realisation. In the items enquiring about the density of wood in relation to water, and sand in relation to water, conservation is not involved at all. Although the performance of the Rintebe group is up to the Grade 6 level (which is 38 per cent.), the level of performance does not suggest that the appreciation of these relative densities develops in the indigenous culture.

Turning now to the comparison between the three sets of Grade 5 performances in Section B of Table 2, a number of observations can be made. Firstly, the performance of the Mt. Hagen group is much superior in the length-displacement group, the conservation of weight, quantity and volume, but not in conservation of area, density ideas, or the effect of immersing objects in water. It is notable that the programme at Mt. Hagen placed heavy emphasis on experiences of measuring lengths, weights and volume, but that the programme included virtually nothing on area, nor did it provide direct experience of different densities. As in the case of weight at Awaba, it seems clear that performance on the test was directly related to educational experience. The poor performance of Siokiei in relation to the Grade 5 total sample is somewhat surprising. The Grade 5 class concerned was a most impressive group. Their achievements in the various Dienes mathematics games was almost dazzling and yet it would seem that this programme, without practical experiences in measuring, had not assisted in developing conservation of the various physical quantities, and this in spite of the fact that it included a good deal of set theory work which might have been expected to develop the general idea of conservation.

V.—ANALYSIS OF RINTEBE TEACHERS' COLLEGE BY TOTAL SCHOOLING AND ENGLISH STANDARD.

The fact that the group at Rintebe Teachers' College contained wide variations in the total schooling of the trainees and also wide variations in the standard of English education attained, and the fact that the standard of English attained bore no particular relationship to the years of total schooling, suggested a possible line of investigation as to the effect of English schooling on conceptual development among the group. It was to be expected that this effect might have been blurred considerably due to the fact that for the most of the trainees English schooling received had been at some time, in some cases a long time, in the past and that any conceptual pattern developed in their English education would have had superimposed back on to it the conceptual pattern of either Pidgin or Kate modes of thought.

TABLE 3

RINTEBE TEACHERS' COLLEGE : PERCENT CORRECT SCORES BY YEARS OF SCHOOLING AND ENGLISH STANDARD.

Topic	No. of items	Total Schooling Years			English Standard			All
		0–3	4–7	8–11	0	1–3	4–6	
Length—Displacement..	4	0	2	10	0	10	10	6
Conservation of Weight	2	67	42	44	45	46	44	45
Conservation of Quantity (plasticine)	3	30	27	23	20	26	31	25
Conservation of Quantity (water)	2	30	49	38	32	48	46	42
Conservation of Area ..	4	0	7	5	5	6	5	5
Conservation of Volume.	2	20	22	17	20	16	21	19
Density Ideas	2	10	24	45	29	32	41	34
Stone in Water	2	89	92	95	86	93	97	93
Wood in Water	2	27	35	33	28	53	37	34
Number in Sample		10	41	60	41	31	39	111

Table 3 shows the performance in the same group test of the Rintebe students, broken firstly according to total years of schooling and secondly by standard of English attained. The results in Table 3 were analysed statistically to determine the probability that they could have occurred by chance. For the test as a whole, the analysis produced the figures in Table 4.

TABLE 4

RINTEBE TEACHERS' COLLEGE : DISTRIBUTION OF SCORES BY TOTAL SCHOOLING AND ENGLISH STANDARD.

	Total Schooling Years			English Standard		
	0–3	4–7	8–11	0	1–3	4–6
Group Mean	43	55	56	49	56	59
Standard Deviation	5	2	2	2	2	2
No. of Students	10	41	60	41	31	39

With regard to total schooling years there is thus no evidence that 8–11 years has a greater effect than 4–7 years. On the other hand, there is a difference between the 0–3 group and the other two, which is significant at the ·01 level. In the case of the English standard attained the differences are significant at the ·001 level.

VI.—DISCUSSION.

We may now review the four lines of evidence which have been presented in relation to the question of how far the conservation of physical quantities is mainly the product of intruding Western education in New Guinea. In the first case we must note that on the science concept test here reported, there was no significant difference in the performance of three most diverse districts of the territory.

The writer understands that this finding is in accord with the results of the primary school final examinations, where the various districts of the territory usually achieve comparable results. The most evident common factor shared by the three districts is the Western style school education system which is working on a common syllabus throughout New Guinea, and which might be expected to produce a common pattern of conceptual development.

Support is given to this view by the evidence of the grade and age comparisons. The fact that, when considered by grade, the performance of the three districts is comparable but that, when considered by age, the performance of the Highland District lags significantly behind the Central Coast, seems clearly established in the present situation. Consideration of the Highlands sample suggests that age is not highly significant in relationships to test performance. It might therefore be inferred that without the intrusion of Western education, the concept of conservation of physical quantities would not be developing in the way revealed by this survey.

In the third place the two sets of comparisons in Table 2 seem clearly to point to the effect of certain particular types of teaching on the development of the concepts concerned. The lagging of Rintebe, where indigenous culture was predominant more than in any of the other groups, the good performance of Awaba Teachers' College in items in which instruction was known to have been given, and most particularly the performance of Mt. Hagen Grade 5 in comparison with other Grade 5s, all point strongly to the effect of Western education. Thus, it seems that particular forms of educational experience associated with measurement must be considered responsible for the performances at Awaba and Mt. Hagen. In this respect, the performance of the Siokiei group was a surprise and somewhat of a disappointment to the writer, as the work in the school had seemed to be directly related to the general concept of conservation. It may be that if the test had measured conservation of number specifically, Siokiei would have performed well. But in this case, if conservation of number was well developed at Siokiei there was little carry-over to conservation of physical quantities generally. It would seem then that not merely Western education, but particularly the quantitative aspects of this, and most specifically experiences in measurement, are responsible for developing conservation of physical quantities in the New Guinea situation.

The fourth line of evidence, that from the closer analysis of the Rintebe results, though by itself by no means conclusive, seems also to point in the same general direction. It is hardly surprising that the small sample of students who had only had up to three years total schooling were overall the group whose physical concepts were most poorly developed. It would be surprising if there were no significant improvement with greater years of schooling, irrespective of medium, especially as this schooling included some mathematical operations and ideas. The lack of improvement over the wide range of education from 4 to 11 years suggests that once a basic level of competence is reached, there is no further improvement with longer vernacular education. On the other hand, there is a significant, though not great, improvement in performance with increasing standard of English education. Restricting ourselves then to the minimum which can be inferred from these results, we may say that English education is more significant than vernacular education in the development of physical concepts. It seems therefore that the Piagetian processes are directly related to Western culture and that the reason that they have been so consistently found in other investigations, is that these investigations have taken place in areas where Western education has to a greater or lesser extent intruded into the indigenous culture. A thorough investigation in an untouched situation would be necessary to confirm this.

The writer's only contribution in this regard comes from the impression of a week-end in a Bena village near the Rintebe Teachers' College, where the only Western intrusion had been the somewhat superficial (from a culture disturbance point of view) influence of the Lutheran Mission, and the fact that two workers from the Summer Institute of Linguistics were resident in the village. As the Lutheran activity had been mainly the teaching of Christianity through interpreters working from Pidgin to the local Bena language, and as Mr. and Mrs. Young of S.I.L. were working entirely in reducing the Bena language to writing and had done only a very limited amount of Bena literacy work, the disturbance was at a minimum. With the help of Mr. Young, some of the science concept test items of the questionnaire referred to above were administered to seven members of the adult literacy class, and a number of them were interviewed subsequently. While quite a number of apparently conserving responses were obtained both in the test and in the subsequent interviews, the proportion of real conservers appeared to be low and there was a high degree of inconsistency about answers given. On the other hand, the attainment of conservation was clearly not simply a matter of degree of European contact. Leaning over the wall of the literacy classroom while the test was in progress was one of the old men of the village, a man highly respected but totally illiterate. While the writer was deforming one of two equal balls of plasticine into a long thin strip, the class and interpreter suddenly dissolved into peals of laughter over a remark from the old man. The interpreter explained subsequently that he had said in Bena, " Don't be fooled by him, he is going to ask you if there is more plasticine but it is only the same piece in another shape." It would be interesting to know whether a high proportion of his own age group had so clear a view, not only of conservation processes, but of what was in the mind of the tester !

The evidence presented in this paper bears only indirectly on the second question put forward in the introduction, that of the relationship of inherited and environmental factors in the conceptual development of children. However, if it is established that Western education is responsible for this particular aspect of conceptual development, then such development must be regarded as being mainly, if not entirely due to environment. There were large differences in the inheritance characteristics of the various groups tested, particularly the Highland sample. The Highland people are now thought to have been descended from a settlement of New Guinea, much older perhaps by an order of several thousand years, than the Melanesian population who came in later and who predominate in the Central and Western districts.

Yet, in spite of such a difference in inheritance, when considered by school grades, there was no significant difference in the performance of the two groups. Further, there were quite a number of mixed race children in the various schools tested in the course of the survey. While these were not analysed separately in detail, it was evident that the performance of the mixed race children corresponded generally to that of the group in which they were found. Those living predominantly in a European manner achieved conservation in much the same way as the Europeans, those living in the village situation performed very much as did their neighbours. Between these two extremes comes a wide range of mixed race homes involving many interdependent variables, the disentanglement of which would be a complex task. The writer has found no evidence, either in the test results of mixed race children or in the observation of particular individuals, that inheritance particularly affects performance in the test administered. On the other hand, there is clear evidence, with the strongest supporting impressions from interviews and all kinds of experiences, that environment is a powerful factor in determining the progress of conceptual development.

VII.—REFERENCES.

GOODNOW, JACQUELINE J. (1962). A test of milieu differences with some of Piaget's tasks. *Psychol. Monogr.*, 76, No. 36 (whole No. 555).

GREENFIELD, P. M. (1966)—BRUNER, J. S., OLVER, R. R., GREENFIELD, P. M., *et al.* *Studies in Cognitive Growth*, Chapter 11. New York: Wiley.

PRICE-WILLIAMS, D. R. (1962). Abstract and concrete modes of classification in a primitive society. *Brit. J. Educ. Psychol.*, 32, 50-62.

PRINCE, J. R. (1967). Science concepts among New Guinea School children—A pilot survey. *Journal of the Papua—New Guinea Society*, Vol. 1, No. 2.

PRINCE, J. R. (1967). Bottlenecks in Papua—New Guinea science teaching. *Education Journal of Papua—New Guinea*, Vol. 5, No. 2.

(*Manuscript received 8th* June, 1967.)

VIII.—APPENDIX.

TEST ITEMS IN THE GROUP QUESTIONNAIRE.

The nature of the items and the type of wording used are illustrated below. The questions have, however, been considerably condensed.

Explanation.	*Question*
1.—Two equal rods are shown to be the same length. One is then displaced lengthwise.	Is the top or bottom one longer or are they both the same length?
2.—A parallel sided board shows a winding and a straight road, both stretching right across the board.	Is the winding road or the straight road longer or are they the same length?
3.—A truck travels along the winding road across the board and an aeroplane travels straight across the top.	Does the aeroplane or the truck go a longer way or do they go the same length of journey?
5.—Two plasticine balls are shown. Both have been made by rolling up three identical sticks of plasticine. Agreement is obtained that the balls are the same size.	Are the two balls the same weight or different weights?
6.—One of the two balls is now rolled into a long thin cylinder.	Is there the same amount of plasticine in the two pieces now, or is there more in the round one or more in the thin one?

10.—Two identical glasses are produced and one nearly filled with water. Water is poured into the other until the group is satisfied that the two glasses contain the same amount of water. The water from one is then poured into an empty glass of wider cross-section, and greater depth.	Is there more water in the thin glass or in the thick glass or do they both have the same amount in them ?
13.—The water level is marked on the glass with chalk.	When the stone is put in the water will the top of the water stay in the same place, go up, or go down ?
16.—Two identical square pieces of cardboard are produced and stated to represent two gardens. Two identical wooden cubes are stated to represent two houses. Agreement to equality of size is obtained. One house is placed near the corner of one garden and the other house in the middle of the other garden.	Which garden has more ground outside the house for growing things, or do they both have the same ?
22.—A plastic tube is produced and bent into a U-shape. One side is marked with black tape and is called the black side, and the other the white side. The group is told that water is to be poured into the white side until the white side is half full.	When the white side is half full will the black side be full right up, empty, half full, or have just a little water in the bottom ?
23.—The tube is going to be tilted so that the bottom moves this way (towards the black side).	Will the water in the tube stay in the same place or move up the white side or move up the black side ?

VERBAL INTELLIGENCE, VISUOMOTOR ABILITY AND SCHOOL ACHIEVEMENT

By M. W. BRENNER and S. GILLMAN

(*Psychological Laboratory, University of Cambridge*)

Summary. The Bender Gestalt Test, scored according to the Koppitz system discriminated clearly between a group of children with visuomotor disorder and educational problems and a matched group of normal subjects.

There was a marked difference between scores for emotional adjustment in the two groups. Failure on the Bender Gestalt Test was associated with under-achievement at school and failure in a selective school examination a year later, even for children with high verbal intelligence. It would appear that selective school examinations are not so heavily biased toward verbal skills as to make success possible for children seriously handicapped by visuomotor dysfunction.

I.—Introduction.

Some concern has been expressed in recent years that the nature of tests used to select children for different types of secondary education is such as to give an unfair advantage to pupils with predominantly verbal skills (Macfarlane Smith, 1964). The present report, together with other studies of visuomotor abilities in school children (Brenner and Gillman, 1966), suggests that pupils with good verbal ability and very poor visuomotor performance are likely to be as severely handicapped as children with the opposite imbalance of skills.

II.—The problem.

A survey of the visuomotor ability of 810 normal children aged 8 to 9 years was undertaken in the schools in the Cambridgeshire area (Brenner and Gillman, 1966). Performance on a group test battery examining visuomotor skills was compared with scores on the N.F.E.R. Primary Verbal Test I. In 8·9 per cent. of children visuomotor performance was superior to verbal intelligence; in 10·2 per cent. of children visuomotor performance was clearly inferior to verbal intelligence. From the latter group a small sample of children, with a marked discrepancy, was selected for further study (the Experimental Group).

The present paper analyses the relationship between verbal I.Q. as measured on two tests, school achievement, and performance on a single individually administered test of visuomotor ability—the Bender Gestalt Test.

III.—Procedure.

The Tests. The Bender Gestalt Test (Bender, 1938) consists of nine designs, presented separately on standard cards; the subject is asked to copy these on a plain sheet of paper. A revised scoring method has recently been standardised by Koppitz (1960), who devised her system especially for screening children. Her score consists of a maximum of twenty-five possible errors, the mean scores and standard deviations for each six months of age from 5 years to 10 years 11 months being based on the records of over 1,000 children. In addition, Koppitz gives ten signs which she regards as significant indicators of emotional disturbance. The Bender Gestalt test in this study was scored according to Koppitz.

The tests of intelligence used were the N.F.E.R. Primary Verbal Test and the Wechsler Intelligence Scale for Children.

The Subjects. The Experimental Group consisted of fourteen children, six boys and eight girls, aged between 9 and 10 years, who were selected for poor visuomotor performance in comparison with verbal I.Q. These children were matched for age, sex, handedness, I.Q., and type of school and home background with fourteen children who showed no appreciable discrepancy between verbal (N.F.E.R.) and visuomotor skills.

IV.—RESULTS.

Table 1 gives the sex, age, I.Q. scores, Bender Gestalt scores and educational assessment of the experimental and control groups.

TABLE 1

SEX, AGE, INTELLIGENCE, BENDER TEST SCORES AND SCHOOL ASSESSMENT OF EXPERIMENTAL (E) AND CONTROL (C) GROUPS.

			I.Q.				
No.	Sex	Age	NFER	WISC verbal	WISC VQ–PQ	Bender Score	School Assessment (Age 8–11)
E1	M*	9·1	115	128	+28	4	Mathematics very poor. Below average generally.
E2	F	9·5	115	115	+21	6	Arithmetic was poor. Handwriting very slow. Keen interest in reading.
E3	F	9·8	112	120	+27	5	Special coaching in arithmetic.
E4	M	9·8	110	128	+28	7	Excessively poor handwriting. Poor spelling and arithmetic. Very keen reader.
E5	M	9·1	106	123	+20	7	Generally below average in all subjects.
E6	M	10·0	104	137	+40	8	Poor written arithmetic. Inability to copy from blackboard. Generally below average.
E7	F	9·2	104	106	+5	1	Average, but handwriting slow and laboured.
E8	F	9·5	100	87	—5	8	C stream. Generally poor.
E9	F*	9·7	95	70	—22	9	Generally retarded.
E10	F	9·9	93	87	+16	8	Promoted from Special Class at 10 years. Considered retarded reader at Secondary School.
E11	M	9·5	90	104	+12	9	Showed promise in Latin at 9 years. Spelling very poor. Arithmetic poor.
E12	F	9·10	90†	109	+ 5	3	Remedial reading at 8 years. Arithmetic below average. Spelling very poor.
E13	M	10·0	91	95	+16	9	Generally below average. Very poor home background.
E14	F	9·11	76†	109	+12	6	Non-reader. Generally below average.
C1	M*	9·0	112	120	—15	1	Above average in all subjects, particuarly Latin and Mathematics.
C2	F	9·3	115	118	—7	0	Above average.
C3	F	9·4	111	121	—28	3	Average. Hardworking.
C4	M	9·7	108	123	—6	0	Above average. Arithmetic good.
C5	M	9·1	110	126	+6	0	Average student. Rather careless worker.
C6	M	9·11	98	111	+14	3	Above average. Near top of form in B stream.
C7	F	9·2	98	104	—3	1	A grade in all subjects.
C8	F	9·6	100	100	—18	0	Above average in all subjects.
C9	F*	9·9	94	111	—2	1	Average. Arithmetic good.
C10	F	9·10	92	94	—3	2	Moved from A to B stream.
C11	M	9·5	96	106	+5	5	Below average. Slow learning to read. Emotional difficulties interfering with school work.
C12	F	9·4	106	106	—3	1	Average. No special problem.
C13	M	9·11	92	105	+5	1	Average. Slow learning to read.
C14	F	9·10	103	97	—2	3	Average. Neat and careful.

* Left-handed children.

† Retarded readers whose controls were matched with their WISC I.Q. Scores.

The mean Bender Gestalt score of the experimental group was 6·4 (standard deviation 2·52). According to Koppitz, this is the mean score to be expected from children of 6½ to 6 years 11 months. The control group showed a mean score of 1·5 (s.d. 1·45), which is very close to the Koppitz scores for this age group (m=1·7, s.d.1·6). The difference between the two groups is significant at the ·01 level.

Two children in the experimental group (E7 and E12) and one in the control group (C11) produced Bender scores which differed appreciably from their performance on the visuomotor test battery. However, in two of these three cases (E12 and C11) the standard of school work was in accord with Bender test score rather than with the group test battery performance. Thus, for twenty-seven of the twenty eight subjects school achievement closely parallelled performance on the Bender test.

Discrepancy between verbal and performance tests on the WISC was not clearly related to Bender Gestalt score. This was due, in part, to the fact that some children (E8 and E9) with very poor visuomotor skills were even more handicapped in verbal skills, and at least one child (C5), with very high visuomotor skill, had even better verbal skill. In the control group, school achievement generally followed verbal I.Q., although obviously motivation and teachers' standards are also involved in both achievement and assessment. In the experimental group, on the other hand, high verbal intelligence, particularly as measured by the WISC, was unrelated to school achievement.

The experimental group showed almost twice as many 'emotional signs' as the control group (confused order: experimental (E) 7 signs, control (C) 2; wavy line: E 2, C 0; second attempt: E 1, C 4; small size: E 2, C 0; expansion: E 0, C 1; overwork: E 3, C 1; total: E 15, C 8). This finding accords with teachers' assessment of the children; with one exception (C11), the children in the control group were considered to be normally adjusted and to have few problems in relationships with staff and other pupils. By contrast, in the experimental group, five children were considered excessively shy and withdrawn, while four were noted for tantrums, aggressive behaviour and unpopularity among staff and follow pupils. It was thought that these personality disturbances were more often the result of failure, frustration and mishandling of these children than the *cause* of their scholastic difficulties.

V.—Discussion and Conclusion.

According to Koppitz (Koppitz, Sullivan, Blythe and Sheldon, 1959), school achievement of first-grade children (aged 6 to 7 years) could be predicted from their performance on the Bender Gestalt test. The exception was the type of child who did relatively poorly on the test, but whose verbal skill and high motivation resulted in good school achievement. In the present sample of children, approximately three years older than the American group, high verbal intelligence and strong motivation did not compensate for poor visuomotor ability, and school achievement closely parallelled Bender Gestalt test scores.

A year after the Bender and WISC tests were completed, these children sat the 11+ Secondary School Selection examination. The results of the examination supported previous school assessment. In the control group, all children with N.F.E.R. I.Q. 108 or above gained places in grammar schools; in the experimental group, only one child (E3) was offered a grammar school place. It appears, therefore, that at this age, the combination of good verbal intelligence with poor visumotor ability results in poor or indifferent scholastic achievement. Evidently, both verbal and perceptual-motor abilities are necessary for academic success at 10 to 11 years.

Acknowledgements.—This work was supported by a grant from The Spastics Society, London. We are indebted to the Chief Education Officer for Cambridgeshire and the Isle of Ely, and to the heads of the schools for their help. The WISC tests were administered by Mrs. M. Farrell and Mrs. P. Koskinen. We wish to express our gratitute to the children and their parents for their co-operation.

VI.—References.

Bender, L. (1938). A Visual motor Gestalt test and its clinical use. *Amer. Orthopsychiat. Ass. Res. Monogr.* 3.

Brenner, M. W., and Gillman, S. (1966). Visuomotor ability in school children. *Dev. Med. and Child Neur.*, 8, 686-703.

Koppitz., E. M., Sullivan, G., Blythe, D., and Sheldon, J. (1959). Prediction of first grade school achievement with the Bender Gestalt Test and human figure drawings. *J. Clin. Psychol.*, 15, 164-168.

Koppitz, E. M. (1960). The Bender Gestalt Test for children—a normative study. *J. Clin. Psychol.*, 16, 432-435.

Smith, I. Macfarlane (1964). *Spatial Ability*. London: Univ. of London Press.

(*Manuscript received 26th November, 1966*)

PERCEPTUAL, REPRESENTATIONAL AND CONCEPTUAL CHARACTERISTICS OF PRIMARY SCHOOL CHILDREN

By DAVID SATTERLY

(*Homerton College, Cambridge*)

Summary. The performances of 200 children aged 7–11 years were investigated in tests of arithmetic, vocabulary and reading, visual and haptic perceptual tasks and drawing and spatial tests. Small groups of children expressing a preference for 'analytic' and 'synthetic' attitudes were identified and their performances compared in perceptual, representational and conceptual variables to investigate the hypothesis of consistency across the operations. Differences between the scores of boys and girls, and older and younger pupils in the entire sample were also studied, and relationships between the variables investigated through correlational and factorial studies.

I.—Introduction.

Studies of the development of perception have shown that, for young children, perception of an item is greatly influenced by the structure of the surrounding field but that this influence decreases with age and gives way to a more analytic appraisal of complex visual stimuli. Piaget (1956) has noted a similar progression in the haptic exploration of shape by children. Literature on children's drawing behaviour suggests that this developmental trend may be discerned in their art products, and Goodenough (1926) saw differences in amount of detail incorporated in children's drawings as explicable in part in terms of analytic or non-analytic perception. Other investigators have employed drawings to study the emergence of concepts of space.

Many attempts have been made to attribute individual differences in perception among adults to persistent preferred methods known as 'attitudes.' Brunswik (1956) has reported a clear distinction between the 'naive' (whole perceiving or synthetic) attitude and the 'analytic' attitude and it has been demonstrated that, whilst it is possible to influence subjects to shift from one mode to the other, different subjects responded to the instructions with varying levels of effectiveness, and some seemed to show a preference for one or other mode of functioning. Though these two modes of perceiving can be clearly distinguished, Vernon (1962) has argued that consistency of attitude in different situations and with different types of material has not been demonstrated and that it is, therefore, not clear whether it is possible to classify perceivers on the basis of their preference.

Witkin and others (1954, 1962) have presented evidence, however, of consistency across a variety of perceptual situations and have proposed a 'global *versus* analytic' mode of approach to designate a cognitive style which expresses itself in both perceptual and intellectual functioning. Gardner and others (1960) have postulated a 'cognitive control' somewhat similar to that of Witkin; and Kagan and others (1964) have investigated analytic and reflective attitudes among children of approximately junior school age range. Some flexibility exists, however, in the concept of analysis and shifts in meaning have occurred in the writings of investigators in this field.

II.—PROCEDURE.

The study reported in this note set out: (i) to investigate the degree of interdependence of perceptual, conceptual and representational operations in middle childhood; and (ii) to discover whether patterning of individual differences in these operations occurs across different situations such that models of personal integration could be applied.

The enquiry was carried out with 200 primary school children (99 boys and 101 girls) with 50 in each of the four age groups (7–8, 8–9, 9–10 and10–11 at the beginning of the school year) in a school on the Bristol fringe of South Gloucestershire. The following tests were administered to all children, tests 3, 7, 11 and 12 being individual tests, the remainder group tests.

(1) Mechanical Arithmetic (Schonell, Form A—revised).
(2) The English Picture Vocabulary Test 2. (Brimer and Dunn, 1963).
(3) Reading and comprehension. The Schonell Graded Reading Vocabulary Test (R 1) and the Silent Reading Test A (R 3) from which a composite score was derived.
(4) Representation of the physical world—drawing test.
(5) The Gottschaldt Simple Figures.
(6) A version of the Gottschaldt Embedded Figures (employed as a test of analytic attitude).
(7) Haptic Perception of Shape—a modification of the test employed by Piaget and Inhelder (1956).
(8) A test of ability to integrate partial visual cues into a whole configuration—measure of synthetic tendency.
(9) A drawing test requiring re-orientation within the three dimensions of space—level of spatial concept.
(10) A test requiring the discrimination between viewpoints—effectiveness of spatial judgment.
(11) Haptic form tracing.
(12) Visual form tracing.
(13) Pictorial depth perception (Hudson, 1960).
(14) Visual discrimination between shapes—speed test.
(15) The Goodenough Draw-a-man Test.

An attempt was first made to identify those children who seemed to express a *preference* for 'analytic' or 'synthetic' perception. Scores on Test 6 and Test 8 were transformed to T-scores in order to render them comparable. The T-score for each child on Test 8 was subtracted from that on Test 6 and a dimension thought of as a continuum ranging from 'analytic' to 'synthetic,' with relatively pure 'types' at the extremes, was constructed. Thus, 'analytics' have comparatively high scores on Test 6 and low scores on Test 8 while 'synthetics' are comparatively high on Test 8, but low on Test 6. At a level of confidence of 5 per cent., it was possible to identify only twenty-four 'analytic' children (fourteen boys, ten girls of mean age 9 : 9; S.D. 12·5 months) and twenty-six "synthetic' children (seventeen boys, nine girls of mean age 9 : 8; S.D. 12·3 months). The 150 children who could not be typed may be thought of as composed of two groups: (*a*) children achieving high scores, on both tests, and (*b*) children having low scores on both tests. Membership of (*a*) implies that the child is effective in whichever mode of functioning the task requires. At the point of split between plus and minus values 94 'analytic' and 106 'synthetic' subjects could be identified. The chi-square value relevant to the distribution of 'analytics' and 'synthetics' by ability stream fell short of significance.

In order to test the hypotheses of consistency of attitude across the operations and in visual and haptic modalities, the performances of 'analytics' and 'synthetics' were compared by analysis of variance to test mean and interaction effects and by the computation of t-ratios in connection with the difference between means. Differences between the mean scores of boys and girls, and of older and younger pupils were also studied.

Product moment coefficients of correlation between scores for each of the four age groups in the sample and for the older (9–11) and younger (7–9) children were calculated in order to investigate the degree of interdependence of the variables and many significant relationships were demonstrated as hypothesised. Latent structure analysis of each matrix was carried out in an attempt to classify the variables, but homogeneity of the data prevented satisfactory grouping. A centroid factor analysis using the technique devised by Thurstone was carried out on the matrix representing the older and younger pupils respectively and a general and two bi-polar factors were extracted from each.

III.—RESULTS.

(1) At each age level tested there is a wide range of individual differences in performance on Test 6 and Test 8. Older children tend to be more analytic than younger children ($p < \cdot 001$), and scores on the analytic-synthetic dimension indicate that there is a progression with age from a preference for a synthetic to preference for an analytic attitude.

(2) Within each sex there are marked individual differences in both analytic and synthetic modes of perception. Nevertheless, in the entire sample boys tend to be more analytic than girls when the task demands it: this difference attains significance ($p < \cdot 01$) among the 9–11 but not among the 7–9 age group.

(3) Preference for analytic attitude does not seem to be related to general school attainment as indicated by the child's presence in A or B stream.

(4) The findings on consistency of attitude do not support the application of models of personal integration. Significant differences in favour of 'analytics' were demonstrated in mechanical arithmetic ($p < \cdot 05$) and in the Gottschaldt Simple Figures ($p < \cdot 01$). 'Analytic' children were, on average, more effective in the picture vocabulary test, drawing, and in pictorial depth perception, in all of which close attention to detail is required. The t-ratios fell just short of the values required for significance at the 5 per cent. level, however. The form tracing experiments, tests 11 and 12, favoured synthetic subjects ($p < \cdot 01$ for test 11, p not significant for test 12). Both require the subject to build up a whole form appraisal following a sequential partial examination. Small but non-significant differences in favour of 'synthetics' were also demonstrated in tests of level of spatial concept and effectiveness of spatial judgment.

(5) Marked sex differences in favour of boys were demonstrated in tests having a spatial component, viz., tests 4, 9, 10 and 13 ($p < \cdot 01$ in all cases among older children and in all except test 9 among the younger children).

(6) The following features of the developmental changes in representing space in drawings were noted:

(*a*) No orderly arrangement of objects (most frequent amongst the younger children).

(*b*) Orderly arrangement in horizontal space achieved but bound to a base line.

(*c*) Some attempt made to represent depth by the employment of the convention of overlapping.

(*d*) Perspective attempted though not always from a consistent point of view.

(*e*) Major problems solved: projective space is depicted, buildings are shown in perspective and distant objects diminish in size.

(7) Two-dimensional responses to the material presented by Hudson(1960) were very common among the younger children. The data support his contention that three-dimensional perception of pictorial material is gradually acquired by children between the ages of 7 and 12 years.

(8) Analysis of performance in the drawing test of level of space concept shows that only few children in the sample had developed a single operational system in which relationships within the three dimensions of space were co-ordinated.

(9) A general factor is seen to be the most important single influence in the factor analyses of performance of older and younger children. None of the well validated factors was identified (with the possible exception of speed and strength of closure among the older group). Similarly, no bipolar contrasting tasks involving analysis with synthesis or the visual with the haptic was extracted. Somewhat different factor patterns are revealed among older and younger children.

TABLE 1

UNROTATED CENTROID FACTORS (100 CHILDREN 9–11 YEARS).

Test	I_0	II_0	III_0	I^2	II^2	III^2	h^2 (obt.)	h^2 (guessed)
1	·667	·281	—·263	·445	·079	·069	·593	·576
2	.728	·175	·227	·530	·031	·052	·613	·682
3	·744	·170	·090	·554	·029	·008	·591	·682
4	·739	—·179	·192	·546	·032	·037	·615	·540
5	·462	·407	·237	·213	·166	·056	·435	·502
6	·633	·236	—·141	·401	·056	·020	·477	·527
7	·683	·124	—·156	·466	·015	·024	·505	·529
8	·643	—·088	—·265	·413	·008	·070	·491	·484
9	·578	—·311	·106	·334	·097	·011	·442	·507
10	·661	·251	·077	·437	·063	·006	·506	·512
11	·334	—·280	—·143	·112	·078	·020	·210	·323
12	·333	—·401	—·146	·111	·161	·021	·293	·401
13	·629	—·108	·307	·396	·012	·094	·502	·540
14	·494	·319	·059	·244	·102	·003	·349	·436
15	·520	—·122	—·110	·270	·015	·012	·297	·432
	Percentage variance			36·48	6·29	3·35	46·12	

TABLE 2

UNROTATED CENTROID FACTORS (100 CHILDREN 7–9 YEARS).

Test	I_0	II_0	III_0	I^2	II^2	III^2	h^2 (obt.)	h^2 (guessed)
1	·611	—·123	—·165	·373	·015	·027	·415	·553
2	·660	—·388	—·160	·436	·151	·026	·613	·680
3	·603	—·422	·225	·364	·178	·051	·593	·680
4	·761	·104	—·329	·579	·011	·108	·698	·704
5	·779	·036	—·119	·607	·001	·014	·622	·623
6	·669	—·239	—·159	·448	·057	·025	·530	·528
7	·598	·361	·147	·358	·130	·022	·510	·537
8	·566	·142	·190	·320	·020	·036	·376	·481
9	·415	—·145	—·171	·172	·021	·029	·222	·363
10	·759	·212	—·060	·576	·045	·004	·625	704
11	·463	·302	·179	·214	·091	·032	·337	·415
12	·385	·249	·285	·148	·062	·081	·291	·381
13	·653	·168	—·264	·426	·028	·070	·524	·571
14	·578	—·130	·249	·334	·017	·062	·413	·429
15	·514	—·078	—·131	·264	·006	·017	·287	·462
	Percentage variance ..			37·46	5·55	4·03	47·04	

Acknowledgements.—The author is grateful to Mr. M. A. Brimer, Head of the Research Unit in the Bristol University Institute of Education, for help given in the preparation of this study, and to Dr. R. Sharp of Hatfield College of Technology for computer facilities with which to calculate the coefficients of correlation.

VI.—References.

Brimer, M. A., and Dunn, L. M. (1963). *Manual for the English Picture Vocabulary Test.* London: N.F.E.R.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments.* Berkeley: Univ. of Calif. Press.

Gardner, R. W., Jackson, D. N., and Messick, S. J. (1960). Personality organisation in cognitive controls and intellectual abilities. *Psychol. Issues, II, No.* 4.

Goodenough, F. L. (1926). *Measurement of Intelligence by Drawings.* New York: World Book Company.

Hudson, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. *J. Soc. Pscyhol.*, 52, 183-208.

Kagan, J., Rosman, B. L., Day, D., Albert, J., and Phillips, W. (1964). Information processing in the child; significance of analytic and reflective attitudes. *Psychol. Monogr.*, 78, 1.

Piaget, J., and Inhelder, E. (1956). *The Child's Conception of Space.* London: Routledge.

Satterly, D. J. H. (1967). *An Investigation into the Relationship between the Two-dimensional Representation of the Physical World by Primary School Children, and Certain Perceptual and Conceptual Characteristics they Exhibit.* Unpub. M. Litt. (Education) dissertation, Univ. of Bristol.

Vernon, M. D. (1962). *The Psychology of Perception.* Harmondsworth: Penguin.

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., and Wapner, S. (1954). *Personality Through Perception.* New York: Harper.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., and Karp, S. A. (1962). *Psychological differentiation: Studies of development.* New York: Wiley.

(*Manuscript received* 13*th March*, 1967)

FLUCTUATIONS IN CHILDREN'S COGNITIVE TEST SCORES OVER A TWO-YEAR PERIOD

By P. S. FREYBERG

(*University of Waikato, New Zealand*)

Summary. Forty-nine New Zealand primary school children were given a group of six tests at three-monthly intervals over a two-year period. Results showed fluctuations in the rate of change of mean score over the nine testing sessions. The mean scores from testing immediately after the Summer vacation showed plateaux in the growth curves, indicating a check in development or performance, both for measures of attainment and intelligence.

I.—Introduction.

It has generally been assumed that children's intellectual development proceeds at a fairly constant rate, at least over such a short period of time as two or three years—barring, of course, any drastic changes in the children's environment or their physiological condition. Fluctuations in test scores have, therefore, been ascribed to idiosyncratic factors and lack of perfect test reliability rather than to any systematic influences.

Fitt (1941) claimed to have identified a seasonal influence on mental functioning, but the evidence he cited is unsatisfactory both for methodological reasons—many of his subjects were observed for less than a year—and because the statistical analyses he made were not always appropriate. However, systematic fluctuations in children's test scores can, indeed, be observed over longer periods of time: in a longitudinal study of children's intellectual development between the ages of 6 and 9 years (Freyberg, 1964), the writer was mainly interested in individual patterns of growth but, nevertheless, examined the changes in mean performance of his experimental group on a variety of tests over a two-year period. The results showed a series of plateaux in the various growth curves, usually coinciding with the school Summer vacations (Figure 1). These plateaux occurred with tests of general intellectual functioning as well as with those of school attainment, and in turn raised questions of the statistical significance of differences in mean increments, and of the implications of this observation as far as standardized testing in schools was concerned.

II.—Description of investigation.

The subjects of the longitudinal study initially comprised 151 children (sixty-four boys, eighty-seven girls) from four schools in a small New Zealand city. They represented the normal range of ability (Primary Mental Abilities I.Q. 77–133, mean 104, American norms). The chronological age range of the sample at initial testing was 5 : 9 to 7 : 10 years, with a median of approximately 6 : 9 years. An experimental group of eighty children from the sample of 151 subjects was then tested at three-monthly intervals over a two-year period (i.e., in August, November, February, May), while a control group, included in the experimental design to enable an estimate of practice effects, was tested only at the beginning and end of this period. Complete longitudinal records were obtained on all the tests for forty-nine children, but for specific tests up to fifty-seven complete records were available.

The tests administered were: Primary Mental Abilities, 5–7 years and 7–11 year forms; Coloured Progressive Matrices; a 72-item multiple-choice test of concept development incorporating Piagetian-type questions; a 23-item dictated word spelling test; a 25-item test of simple arithmetical problems; and a 120-item speeded test of arithmetic computation, involving sub-tests of the four basic processes. All except the first two of these instruments had been constructed and pre-tested by the writer, and yielded reasonably satisfactory coefficients of reliability (retest after one week, minimum ·82 for single year age group, with median $r = \cdot 91$). The various tests were introduced into the experimental series at the appropriate levels as noted in Figure 1, and records were thus available for at least twelve months (five test periods) in each case. Significant practice effects were observed in the case of the experimental group for all except the two arithmetic tests.

III.—Results.

A number of problems arise when one attempts a comparison of the mean performance of the group at each test period. First, the effects of practice or familiarity may obscure changes in mean level of performance attributable to maturation and other factors; and it has generally been observed that practice effects with intelligence tests are greatest at the second and third administrations and diminish rapidly thereafter (McIntosh, 1944; Wiseman and Wrigley, 1953). Second, in attainment testing the difficulty level of test items may alter substantially as new material is taught during the school year. Third, it is not sufficient just to observe any fluctuations in the curves of mean performance: the significance of such fluctuations (if any) must be established.

If all the mean performance curves were of a similar nature, it would be possible to estimate the probability of any departures from regularity, but with the variety of tests used, and differences in extent of practice effects, such similarity in mean performance curves is hardly to be expected. Thus, the alternative had to be adopted of examining each child's increments on each test between one period and the next. This made possible an analysis of variance of increments, testing variation among mean increments between successive periods against variation between subjects and the subject-period interaction. The variance ratio was significant for all six tests, p being at least $< \cdot 01$ in each case.

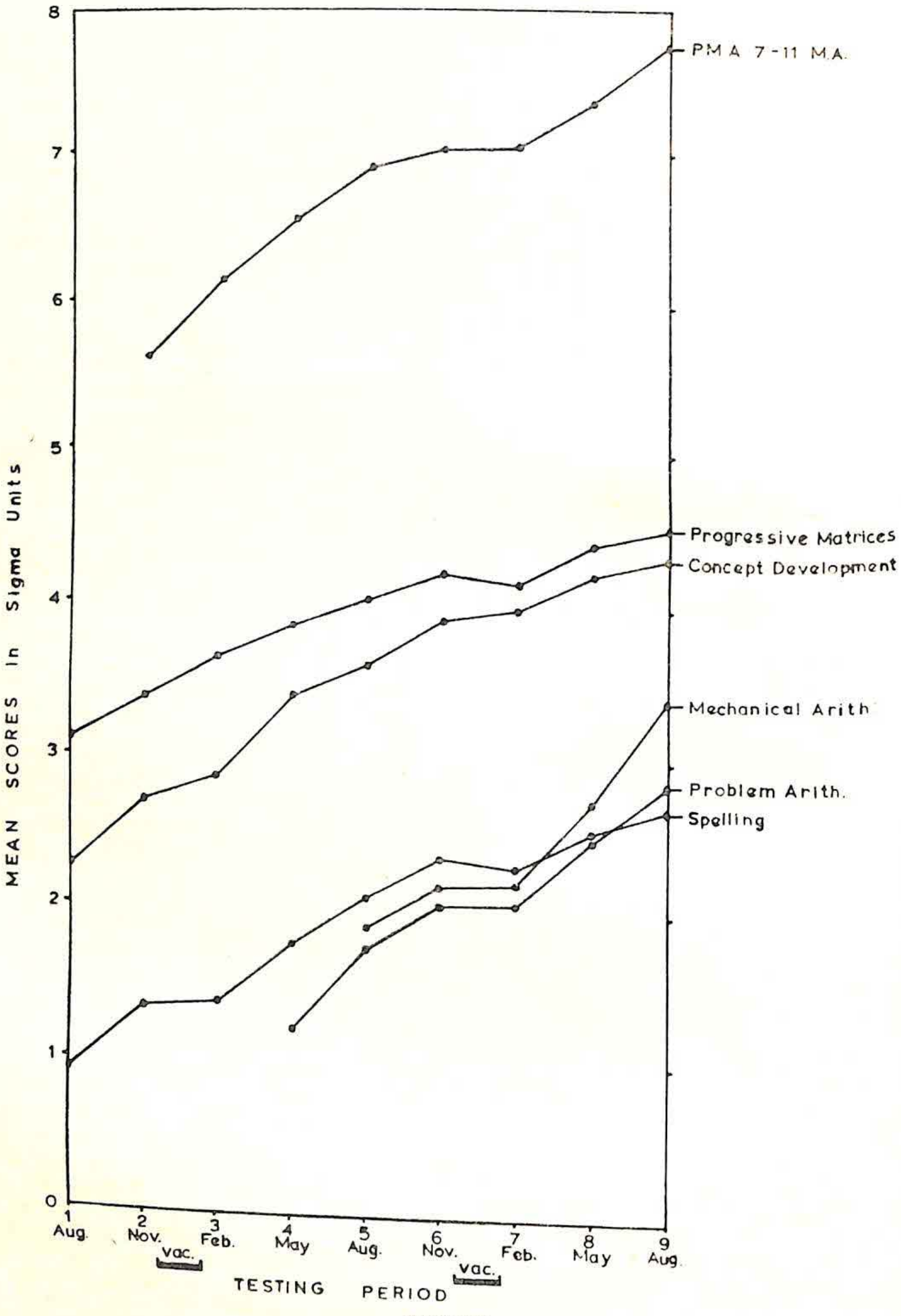


FIGURE 1

Graph of mean scores for complete longitudinal group at each testing period (N=49).

The differences between mean increments were then examined. It is, of course, not legitimate statistically to select arbitrarily the largest mean increment and assess the significance of its difference from the smallest increment by means of an ordinary t test, as Fitt (1941) did. What is required is some method which enables an estimation of the significance of differences among all increments of the observed sizes. The further the separation in ranking between the various mean increments, the larger the t ratio necessary to establish the significance of the difference, because the data have been ordered. The problems associated with this type of analysis have been discussed in detail by Duncan (1955).

The solution suggested by Duncan is a multiple-range t test of ordered means which takes into account the positions of the compared means in the ordering. An examination of the test results using this technique brought out the following points :

(*a*) For the two arithmetic attainment tests, mean 6 (i.e., the mean increment between the sixth and seventh testing periods, November-February) is significantly smaller than all other mean increments on these tests. The February testing, it should be noted, was undertaken in the third week of the new school year.

(*b*) For the spelling test, mean 6 is significantly different from all other means except mean 2 (another November-February interval). Mean 2 in this case is not significantly different from mean 8, but at this stage of the test programme a ceiling effect with this test was observed.

(*c*) For the PMA 7–11 test, mean 6 is significantly different from all other means except mean 5. Mean increment 2 does not result in a plateau as in this case it seems to include the substantial practice effect usually observed between the first and second times a test is administered.

(*d*) In the case of Coloured Progressive Matrices, mean increment 6 is by far the smallest—actually it is negative—but the difference between it and the other increments does not quite reach the 5 per cent. level of significance of difference. Practice effects again seem to have obscured the plateau expected at mean increment 2.

(*e*) In the case of the concept development test, means 6 and 2 are again amongst the smallest, but are not significantly different from the remaining mean increments.

IV.—DISCUSSION.

While the results of the above analysis are not as clear-cut for the tests of general intellectual abilities as they are in the case of school attainment, they suggest that assumption of constancy in rate of change of *measured* intelligence over even short periods of time is not warranted. Whether the fluctuations in mean scores are caused by a seasonal factor or by the long Summer vacation cannot be determined from this investigation : a similar study where children were absent from school for several weeks at another season of the year would be necessary to answer this question.

The most likely explanation of the plateaux in score curves on all the tests administered here is simply that the children's performance is temporarily lowered by lack of continued practice in the tackling of more formal intellectual tasks. Most psychologists now agree that so-called intelligence tests can no longer be regarded as testing innate capacity or even capacity itself, and it is therefore not surprising that such tests should reflect recent changes in environmental experience in the same way as attainment tests do, though possibly to a lesser extent. What is important is to realise the effects of this on intelligence test scores that are obtained from testing within the first month or so after the Summer vacation.

The publishers of some reputable attainment tests have adopted the practice of quoting three sets of norms—for tests administered at the beginning, in the middle and at the end of the school year. This has reduced, to some extent, the possibility of teachers recording spuriously large changes in performance level for their children from tests taken at the start and finish of the year. But if the ' beginning ' norms have been derived from interpolation rather than separate administration, the present evidence indicates that they may over-estimate the mean level of performance during the first month, whether percentile ranks or class equivalent transformations are used. Similarly, the ' average ' 9-year-old tested immediately after the long vacation may well score lower on a test of general intellectual ability than the ' average ' 9-year-old tested in the middle of the school year. Interpolated means for the beginning of the year for the attainment tests described above would have been inflated by approximately 13 to 25 per cent. of the full year's increment when compared with the observed means for the various tests—a difference sufficiently large to be not only of statistical but also of considerable practical significance.

Replication of the present investigation is obviously necessary, to increase the generality of the sample, both in representativeness and in age range. Should confirmation of the results be obtained, it is also obvious that one of two alternatives must be seriously considered : either to avoid the administration of intelligence and attainment tests in, say, the first six weeks of the new school year, or to provide separate (and not merely pro-rated) norms for tests given during this period.

V.—REFERENCES.

DUNCAN, D. B. (1955). Multiple range and multiple *F* tests. *Biometrics*, 11, 1-42.

FITT, A. B. (1941). *Seasonal Influence on Growth, Function and Inheritance.* Wellington : New Zealand Council for Educational Research.

FREYBERG, P. S. (1964). *Some Aspects of Intellectual Development in Children Aged 6 to 9 Years.* Unpublished Ph.D. thesis, Victoria University of Wellington.

MCINTOSH, D. M. (1944). The effect of practice in intelligence test results. *Brit. J. Educ. Psychol.*, 14, 44-45.

WISEMAN, S., and WRIGLEY, J. (1953). The comparative effects of coaching and practice on the results of verbal intelligence tests. *Brit. J. Educ. Psychol.*, 44, 83-94.

(*Manuscript received 26th July, 1967*)

VARIATIONS IN DRAWINGS OF ' A PERSON ' and ' MYSELF ' BY HEARING-IMPAIRED AND NORMAL CHILDREN

BY JOHN GILLIES

(*Department of Psychology, University of Glasgow*)

SUMMARY. Study of drawings of ' a person ' and ' myself ' by hearing-impaired children suggested the existence of consistent patterns of qualitative difference between the two. A simple method of quantitative comparison was devised and comparisons made with drawings by a control group of normally hearing children. Analysis of results showed significant variations between the hearing-impaired and normal groups in terms of direction and degree of person/self differences for the two types of drawing. These results, together with those relating to possible sex differences, are discussed and a wider investigation proposed to consider the possible relevance of these trends to conceptual development and psycho-social adjustment.

I.—INTRODUCTION.

The following observations arose in the course of a larger study, projected in 1963 (Bowyer, L. R., Marshall, A., and Wedell, K., *Brit. J. Educ. Psychol.*, 33, 85-87, 1963), and at present being carried out in the University of Glasgow under a grant from the Scottish Council for Research in Education, into the relative social-emotional adjustment of severely deaf and partially-hearing children in ordinary and special schools. Among the variety of reportive, observational and projective data obtained in respect of each child were drawings of ' a person ' and ' myself.'

These drawings were examined by the investigators not according to any strict set of criteria but rather as suggestive supplements to more comprehensive projective assessments. In the course of this examination, however, the impression grew that certain definite patterns of difference between ' self ' and ' person ' drawings were discernible. It appeared that drawings of self among the deaf children tended, on the whole, to be larger and richer in detail than drawings of a person, while from a small number of such drawings then available, the reverse seemed true in the case of normally-hearing children. In order to provide a general check on this possibility, the present short study was devised, using all suitable drawings by the deaf available from the larger study, together with a comparable sample of drawings by normally-hearing school children.

II.—SUBJECTS.

A.—Deaf/Partially-hearing Group.

The subjects in this group were the forty-nine severely deaf and sixty partially-hearing children for whom testing had been completed in the main study referred to. Of these, the forty-nine severely deaf and twenty of the partially-hearing were pupils in special schools for the deaf in the Glasgow area, while the remaining forty partially-hearing children were attending ordinary day schools in the city. The age range in both sub-groups was of necessity fairly wide (7–15 years in the severely deaf, 5–15 in the partially-hearing group) since the subjects included most of the pupils in schools for the deaf in the area.

B.—Normally-hearing Group.

The 105 subjects comprising the control group were children of normal hearing attending conveniently situated schools in the Glasgow area. Drawings were obtained via student teachers studying in the Psychology Department of Glasgow University.

Some attempt was made to match the control group to the deaf/partially-hearing group in terms of proportions of males to females and of age ranges, but in both cases matching could be no more than approximate.

Before final scoring, some twenty-one subjects, mainly in the hearing-impaired group, had to be eliminated since their drawings were not suitable for scoring, either because they were manifestly not representations of the desired subjects or were so idiosyncratic as to be unadaptable to the scoring method (q.v.). The final distribution of subjects was: deaf group, seventeen boys, twenty girls; partially-hearing, twenty-seven boys, twenty-four girls; normally-hearing, forty boys, sixty-five girls.

III.—PROCEDURE.

Collection of drawings.

The procedure employed in obtaining drawings from the hearing-impaired children was as follows: The child was presented with a sheet of drawing paper of a standard size (10-in. × 7-in.) and a pencil, with the instruction, " Draw a person." Where the term ' person ' proved to be difficult, the phrase ' a boy or a girl ' was substituted and the request was reinforced if necessary by writing or pantomime. No time limit was set for the task and no attempt made to influence the quality of the product. When the child indicated that he had finished, the sheet was turned over and the instruction " Now draw you " given. The endeavour was always to keep conditions as similar as possible for the production of the two drawings.

The drawings from the hearing group were obtained in a similar manner with the co-operation of school teachers.

Scoring.

The drawings of each subject were scored on the original scale of the Goodenough Draw-a-Man Test (1926) in order to obtain a simple quantitative assessment based on developmentally-linked features. The raw scores in each case (' person ' and ' self ') were compared and the following factors considered:

A. Direction of shift from person score to self score, scored + if the self score is higher, − if the self score is lower, O if both scores are the same.

B. Degree of shift, i.e., amount of difference between self and person scores.

C. Size differences between the two drawings, scored + if self larger by half than person, − if person larger by half than self, O if difference is less.

IV.—RESULTS.

The results of this analysis of the drawings are set out in Table 1. Appropriate significance tests (X^2 and difference of means) applied to the three variables yield the results summarised below. A break-down of the deaf group into severely deaf and partially-hearing (males and females) revealed no significant differences between these sub-groups.

TABLE 1

COMPARISON OF DRAWINGS OF 'PERSON' AND 'SELF.'

	Deaf			Hearing		
	Girls	Boys	All	Girls	Boys	All
Direction of Shift:						
No. of cases +	23	25	48	20	13	33
No. of cases O	6	8	14	12	5	17
No. of cases −	15	11	26	33	22	55
Degree of Shift:						
Mean + shift	5·26	4·64	4·94	2·10	2·77	2·36
Mean − shift	3·67	3·73	3·69	2·70	3·59	3·05
Mean shift ±	4·00	3·52	3·78	2·02	2·88	2·34
Difference in Size:						
No. of cases +	21	14	35	11	11	22
No. of cases O	19	18	37	24	14	38
No. of cases −	4	12	16	30	15	45

+ = 'Self' score higher or 'Self' drawing larger.

A. *Direction of shift.*

The deaf group showed predominantly positive shift, the hearing predominantly negative shift ($p < \cdot 01$).

B. *Degree of shift.*

The deaf group showed a larger mean shift (positive or negative) than the hearing group ($p < \cdot 01$), but whereas deaf positive shifters showed a greater mean shift than hearing positive shifters ($p < \cdot 01$), deaf negative shifters did not show a significantly higher mean shift than hearing negative shifters. It is also of note that while female deaf subjects showed significantly greater mean shift than hearing females, irrespective of direction ($p < \cdot 01$), there was no significant difference between deaf and hearing males.

C.—*Size.*

The deaf 'self' was predominantly larger, the hearing 'self' smaller than the 'person' ($p < \cdot 05$). Again, the female deaf differed highly from the female hearing ($p < \cdot 01$), whereas males showed no significant difference.

Despite the encouraging levels of statistical significance on the major variables, it must be admitted that experimental controls were rudimentary and criteria somewhat arbitrarily adopted. It is proposed to replicate the investigation on a larger scale using stricter experimental techniques of matching and control, and possibly a refined version of the scoring procedure which might allow (*a*) more variables to be assessed and (*b*) the inter-relations of the scored factors to be clarified. Should the trends suggested by the present study be confirmed, it may be possible to go on to test various interpretations. Do the discrepancies between the deaf and hearing groups relate, for example, to variations in cognitive development (e.g., body-image distortion), differential operation of practice or fatigue factors in the two groups, or psycho-social adjustment? When the large-scale adjustment study of the present subjects has been completed, it may be possible to determine whether degree of adjustment relates significantly to any of the patterns of variation observed here for the hearing-impaired group. A similar investigation of the control subjects should be possible. It should also be possible to devise means of investigating the relation of these variations to differences in body-image attainment and to practice or fatigue factors, and in addition, to reach an interpretation of the apparent sex differences in score patterns. (The *female* deaf and hearing groups differ significantly ($p < \cdot 01$) on both mean shift and size change, while there are no significant differences between deaf and hearing *males* on the corresponding factors, nor are there significant sex differences *within* either group).

(*Manuscript received* 13*th July*, 1967)

CHARACTERISTICS OF REFERRALS BY SCHOOL PSYCHOLOGISTS AND GENERAL PRACTITIONERS TO A CHILD PSYCHIATRY CLINIC

By JOHN C. PETRICCIANI
(*Department of Pediatrics, Stanford University, U.S.A.*)

AND SULA WOLFF
(*Department of Psychological Medicine, Royal Hospital for Sick Children, Edinburgh*)

SUMMARY. A preliminary comparison was made of several major characteristics of 140 children referred to a child psychiatry clinic by general practitioners and school psychologists in Edinburgh. Over a twenty-eight-month period three times as many children were referred by general practitioners as by school psychologists. General practitioner referrals had normally distributed intelligence test scores while those of school psychologist referrals were skewed towards the below average category. Children from the most affluent area of Edinburgh were over represented among general practitioner referrals, while children from the area with the most pronounced social problems came primarily from school psychologists. The attrition rate for the 140 children was about 25 per cent. ; the initial contact with social workers may have been important for this group.

I.—INTRODUCTION.

Most children with psychological problems are identified as disturbed either by their parents or by their teachers. The first group of children reach the psychiatric clinic via the family doctor ; the second via the school psychologist. The purpose of this study was to make a preliminary evaluation of the two referral groups with respect to their composition in terms of age and sex, the type of problems which precipitated the referral, social variables, and the interaction of the family with the clinic.

II.—METHOD OF INVESTIGATION.

(*a*) *Selection of Cases.*

Seventy consecutive referrals by Edinburgh school psychologists were chosen from the files of the Department of Psychological Medicine at the Royal Hospital for Sick Children in Edinburgh. The period covered by these seventy cases was twenty-eight months, during which time there were 229 general practitioner referrals of children in the same age group (5 to 15 years). Seventy of the 229 general practitioner referrals were selected using a table of random numbers.

(*b*) *Classification of Cases.*

The child's presenting symptoms were classified using Rutter's (1965) diagnostic scheme. Each case was denoted as one of the following: (*a*) neurotic (enuresis, encopresis, fears, stammering, school refusal, disturbed sleep, tics, crying, and nervous mannerisms); (*b*) conduct disorder (aggressive behavior, stealing, lying, truanting, temper tantrums, attempted suicide, sexual acting-out, and generally difficult behaviour or failure to conform); (*c*) mixed (*a*) and (*b*) if both were a major feature of the case; (*d*) other (learning problems, withdrawal or autism, somatic complaints, restlessness and depression). The internal consistency of the system was demonstrated by an 80 per cent. agreement when the authors each independently classified a sample group of cases.

(*c*) *Intelligence.*

The intelligence of sixty-nine out of seventy of the school psychologists' referrals had been assessed; only fifty-two out of seventy of the general practitioners' referrals had been tested. Individual test scores were used, based on either WISC of Stanford-Binet. Each case for which information was available was categorized as either below average (less than 90), average (90–110), or above average (greater than 110).

(*d*) *Social Variables.*

The data available in the case notes were insufficient to classify the families according to social class. Home addresses permitted a rough classification using the material recently published by Philip and McCulloch (1966), who related eighteen social indices to the twenty-three electoral wards in Edinburgh. The distribution of cases over the twenty-three electoral wards for each referral group was used to calculate the rate of referral for every 1,000 children at risk. On the basis of these rates, the electoral wards were ranked. A comparison of the rank order for each referral group was then made together with appropriate correlations with the social variables described by Philip and McCulloch.

(*e*) *Interaction with the Clinic.*

The following indices of the family-clinic interaction were selected because of the accessibility of quantitative data: the number of home visits attempted and completed; the number of first contacts with the family by a psychiatric social worker, student psychiatric social worker, or psychiatrist; the number of failed first clinic appointments; and the number of cases closed because of poor family co-operation.

III.—STATEMENT OF RESULTS.

Table 1 shows that there is no difference in the distribution of symptoms by sex between the two referral groups when our broad categories were used. Two specific diagnostic categories (brain damage and school refusal) were, however, differently distributed between the two referral groups; two children with minimal brain damage came from general practitioners as opposed to seven from school psychologists, and only three school refusals from general practitioners compared to eleven from school psychologists.

Table 2 shows the age distribution for the two groups. The mean age of the school psychologist referrals is 9·7 years which is not statistically different from 8·5 years in the general practitioner group. There is, however, a distinguishing peak of referrals in the school psychologist group at the age of 13, which is largely (7/12) due to school refusals and corresponds to the year in which most students transfer from primary to secondary school. The female patients tended to concentrate in the 8 to 11-year range in both groups and there was little difference in the number of female patients in each group. More children are referred between the ages of 5 and 8 years by the general practitioners (twenty-five) than by the school psychologists (ten). With the exception of the 13-year peak in the school psychologist group, there is an overall decline in the number of children referred after the age of 9 or 10.

TABLE 1

SET DISTRIBUTION OF SYMPTOMS FOR EACH REFERRAL GROUP.

	School Psychologist			General Practitioner		
	Male	Female	Total	Male	Female	Total
Neurotic	16	5	21	13	8	21
Conduct Disorder	26	6	32	24	10	34
Mixed	4	—	4	1	—	1
Other	9	4	13	12	2	14
All Symptoms	55	15	70	50	20	70

TABLE 2

AGE DISTRUBUTION FOR EACH REFERRAL GROUP.

Age		5	6	7	8	9	10	11	12	13	14	15	Total
School psychologist	M	1	3	4	10	5	11	3	4	11	2	1	55
	F	0	1	1	3	3	1	3	0	1	1	1	15
General Practitioner	M	6	9	9	8	8	1	5	0	3	1	0	50
	F	0	0	1	5	4	5	1	1	1	2	0	20

The intelligence test scores of children referred by general practitioners were normally distributed (eight below average; thirty-one average; thirteen above average). The distribution of test scores of children referred by school psychologists was skewed towards the below-average (thirty below average; twenty-nine average; ten above average).

There was no overall statistical difference between the two groups in the rates of referral from the various electoral wards in Edinburgh. The referral rate (per 1,000 children in the 5 to 15-year range) for each of the wards fell within 1·6 standard deviations of its respective mean except the ward having the highest referral rate in each group, and both of these were three standard deviations above their mean (3·05 and 3·56). Since these two wards were so easily distinguishable from lower ranking wards, the social variables with which they correlated highly are presented below. The ward with the highest rate of school psychologist referrals (St. Giles) also had the highest rate of attempted suicide; the second highest of juvenile delinquency, of children being taken into care, and of RSSPCC referrals; and the third highest of school absences other than for sickness. The referral rate from general practitioners was highest from the ward (Morningside) which also ranked first in the number of owner-occupied houses; last in the number of children taken into care and overcrowding; and last but one in juvenile delinquency.

Table 3 shows that cases referred by school psychologists were less often assigned to social worker students than general practitioner referrals, and that in the former group the initial contact was more often with the psychiatrist himself. School referrals usually came to the clinic with more detailed information available and requests for a diagnosis only were more common in the school psychologist group (nine) than in the general practitioner group (one). There was no difference between the two groups in the number of home visits attempted or completed, which indicated that the clinic expended an equal effort in obtaining social data from each group. If the number of failed first appointments at the clinic is taken as a measure of the initial reluctance to enter into therapy, it can be seen that there was no significant difference between the two groups. Approximately equal numbers of cases in each group were closed because of poor family co-operation.

TABLE 3

INDICES OF THE FAMILY-CLINIC INTERACTION.

	School Psychologist n=70	General Practitioner n=70
First Contract :		
PSW	23	18
PSW Student	23	37
Psychiatrist	19	10
Unknown	5	5
Home Visits Attempted	44	48
Home Visits Completed	43	44
Failure to Keep First Clinic Appointment	16	13
Cases Closed due to Poor Co-operation	18	19

Overall about one-quarter of the cases were closed because of poor co-operation on the part of the patient or parent(s). The distribution of symptoms in this sub-group was the same as that in Table 1. The unco-operative cases in the school psychologist group dropped out more rapidly than those in the general practitioner group, with only one of eighteen patients in the former returning to the clinic after the first visit but six of nineteen in the latter returning for more than four appointments. In the cases closed because of poor co-operation, the initial contact with the parents was by a psychiatric social worker or student psychiatric social worker in all but one, and these cases were divided equally between the professional and student social workers.

IV.—DISCUSSION AND CONCLUSIONS.

In Scotland the school psychological service conducts its own child guidance clinic separate from the psychiatric service for children provided under the National Health Service. Only a small minority of children seen by educational psychologists are referred to child psychiatrists. At the Psychiatric Department of the Edinburgh Children's Hospital, three times as many children are referred by general practitioners as by school psychologists. This is in contrast to a recent report by M. Hare (1966) in which there were approximately equal numbers of general practitioner and school medical officer referrals to the Croydon Child Guidance Clinic in Surrey.

The difference in referral rates between the two referral groups in the age range 5 to 8 years may represent a lower tolerance of the children's symptoms by the parents than by the teachers, or it may be due to a more successful handling of this age group by the school psychologists, so that while the children were recognized by their teachers as having problems, fewer were referred to our clinic because of adequate treatment by the school psychologists. It is interesting to note that the significant decline in the number of children referred by both general practitioners and school psychologists begins just before adolescence and falls to low levels at a period in which one would expect the prevalence of psychiatric disorders to be high.

Even though there was no overall distinction between the two referral groups in terms of the referral rates from the various wards within the city of Edinburgh, the two wards having the highest rates were easily distinguishable and represented rather opposite segments of the community with respect to the social variables considered. Thus, on a *per capita* basis, what might be termed the most affluent area of the city makes the greatest use of the clinic through general practitioner referrals while children from the area with the most pronounced social problems come to the clinic primarily through school psychologists. This finding agrees with earlier work in the United States by Pfouts (1963) and more recently Furman (1965) who showed that the route by which patients come to psychiatrists is directly related to their socioeconomic status, with referral by public agencies being correlated highly with low social class while referrals from other medical facilities were frequently patients of higher income.

The extent of poor co-operation on the part of the patient or parent(s) in this study was considerably lower than the 0·59 attrition rate reported by Tuckman (1959) or the 0·57 in Overall's (1963) study. The more rapid attrition rate in the school psychologist group may be due to the fact that in general they have already had a number of school clinic visits when they first come to the child psychiatry clinic. In view of the fact that equal numbers of patients in this subgroup were seen by PSW's and student PSW's, it seems likely that poor co-operation is related to the function of the social worker rather than the extent of her experience. In this regard one might postulate the existence of a number of patients (parents) in whom the initial contact is of great importance with respect to continuance in therapy. Recent studies by Chance (1959), Heine (1960), and Lennard (1960) have shown that if a patient's expectations of the role which his initial contact should assume go unfulfilled and neither the patient nor the clinic contact is willing to alter his position, there will be great difficulty in establishing a working relationship and continuing in therapy. Thus, the parent's expectations of the social worker may be quite different from the collection of a detailed social history which is almost always the social worker's function on the first interview. In as much as we are unable to identify those parents who are at greatest risk with respect to the conduct of the initial interview, perhaps the best approach towards reducing the number of cases terminated because of poor co-operation is to simply pay particular attention to the role expectations of the parents and to attempt to reconcile any large differences between their expectations and reality before the initial interview is concluded.

ACKNOWLEDGEMENT.—This study was supported in part through United States Public Health Training Grant 8 TI-HD-49-07. We are grateful to Dr. M. M. Methven of the Department of Psychological Medicine at the Royal Hospital for Sick Children, for her co-operation in making the case records available. Thanks are also due to several members of the Departments of Psychiatry at the University of Edinburgh and Stanford University for their criticism and helpful advice.

VI.—REFERENCES.

CHANCE, E. (1959). *Families in Treatment*. New York: Basic Books.

FURMAN, S., *et al.* (1965). Social class factors in the flow of children to outpatient psychiatric facilities. *Am. J. Pub. Health*, 55, 385-92.

HARE, M. K. (1966). Short treatment in a child guidance clinic. *Brit. J. Psychiat.*, 112, 613-6.

HEINE, R. W., and TROSMAN, H. (1960). Initial expectations of the doctor-patient interaction as a factor in continuance in psychotherapy. *Psychiat.*, 23, 275-8.

LENNARD, H. L., and BERNSTEIN, A. (1960). *The Anatomy of Psychotherapy*. New York: Columbia University Press.

OVERALL, B., and ARONSON, H. (1963). Expectations of psychotherapy in patients of lower socioeconomic class. *Am. J. Orthopsych.*, 33, 421-30.

PFOUTS, J. H., *et al.* (1963). An outcome study of referral to psychiatric clinic. *Social Work.*, 83, 79-86.

PHILIP, A. E., and MCCULLOCH, J. W. (1966). Use of social indices in psychiatric epidemiology. *Brit. J. Prev. Soc. Med.*, 20, 122-6.
RUTTER, M. (1965). Classification and categorization in child psychiatry. *J. Child Psychol. Psychiat.*, 6, 71-83.
TUCKMAN, J., and LAVEL, M. (1959). Attrition in psychiatric clinics for children. *Public Health Reports*, 74, 309-15.

(*Manuscript received 6th September*, 1967)

TEACHER RATINGS OF EXTRAVERSION AND NEUROTICISM AND CHILDREN'S INVENTORY RESPONSES

BY S. B. G. EYSENCK
(*Department of Psychology, Institute of Psychiatry, London, S.E.*5)

AND A. J. PICKUP
(*Bromley, Kent*)

I.—INTRODUCTION.

The purpose of this study was to compare the questionnaire responses of boys and girls on the Junior Eysenck Personality Inventory (S. B. G. Eysenck, 1965) with ratings furnished by their teachers. It was not anticipated that the results would throw much light on the problem of validity of the questionnaire, as teacher ratings are suspected of being governed to a large extent by halo factors of various kinds (Eysenck, 1960); also previous work with normal adults had shown that the neuroticism-stability variable is difficult to rate because the factor is concerned largely with internal feelings accessible to introspection, but not to external scrutiny (Eysenck and Eysenck, 1968). It was thought, however, that the results would throw some light on the nature and existence of the halo factor itself, and that some relationships might emerge between ratings and inventory responses for extraversion-introversion, which had previously been found to be easier to rate than neuroticism.

II.—PROCEDURE AND RESULTS.

Ten secondary schools in two London suburbs were selected and furnished nineteen classes representing a wide range of age and ability. The majority of these classes were second and third year pupils of average and below average ability.

Eighteen teachers of these classes participated, representing a wide range of subject specialization and teaching experience. The majority of these teachers were English and Mathematics specialists and most had taught for between ten and fifteen years. The experimenter visited the school during a period when the participating class was normally taken by the participating teacher. The teacher would complete his estimates in private, while the experimenter would administer the J.E.P.I. to the children under conditions of anonymity. Teachers were supplied with a pen-picture of the traits characterizing extraverted and introverted personalities and were asked to rate each pupil on four bi-polar scales; Extraversion, Stability, Maturity and Co-operation. These scales were constructed to offer no neutral mid-point, as it was feared that most children would have been rated at this point. The teacher was thus forced to rate each child as (i) mildly or (ii) extremely characterized by each trait, or its opposite (e.g., Stable . . . Unstable). The results of the experiment are shown in Table 1, for boys and girls separately. E (inventory scores on extraversion) correlate significantly with Ex (extraversion ratings) for most groups (r= ·24 for all 500 pupils); these correlations show a tendency to increase with the age of the pupils ; and the correlations for girls are conspicuously and significantly higher than those for boys. N (inventory scores for neuroticism) correlate only very little with St (teachers' ratings of stability). The correlation over all 500 pupils is ·08, which is not significant ; note that the correlation is positive between Neuroticism as experienced and Stability as rated !

TABLE 1

INVENTORY SCORES (E AND N) AND TEACHERS' RATINGS (EX AND ST).

	Boys			Girls		
Age	N	E×Ex	N×St	N	E×Ex	N×St
11	38	·25	·24	—	—	—
12	64	·06	·13	75	·23*	·16
13	91	·21*	—·01	79	·23*	·30**
14	73	·32**	·03	62	·49***	—·10
15	18	·32	·15	—	—	—
Total	284	·21***	·07	216	·31***	·12

* p< ·05. ** p< ·01. ***p< ·001.

The reason for this failure of teachers and self-ratings on neuroticism to agree becomes apparent from the results shown in Table 2 ; here C stands for teachers' ratings on co-operation, M for ratings on maturity. It is clear that teachers have a generalized halo concept of co-operation, stability, maturity and, to a much smaller extent, extraversion ; in other words, whatever socially desirable trait they are asked to rate, they judge the child in terms of his agreeableness in the school situation (co-operation). This makes it unlikely that this halo concept should correlate with the child's self-perceived neuroticism, and in so far as these neurotic troubles make the child less noisy and troublesome, they may in fact cause the teacher to rate him as more stable !

TABLE 2

INTERCORRELATIONS OF TEACHERS' RATINGS.

	Boys				Girls			
Age	N	C×Ex	C×St	C×M	N	C×Ex	C×St	C×M
11	38	—·16	·53***	·60***	—	—	—	—
12	64	—·09	·29**	·32*	75	·09	·46***	·26*
13	91	·45***	·71***	·54***	79	·11	·47***	·14
14	73	·35**	·52***	·56***	62	·28*	·72***	·31*
15	18	·63**	·58*	·23	—	—	—	—
Total	284	·18**	·55***	·51***	216	·16*	·60***	·25***

* p< ·05. ** p< ·01. *** p< ·001

The Lie scale of the J.E.P.I. failed to show any significant correlations with the ratings. It had been thought that possibly L might correlate negatively with maturity, but the overall correlation was —·02, and the correlation failed to be significant or even suggestive in any of the different age and sex groups.

III.—CONCLUSION.

The results altogether do not suggest that teachers' ratings, undertaken under conditions such as those described, furnish us with results which show any close relationship to children's personality as revealed in self-ratings. Whether teachers' ratings or self-ratings are more valid, and give higher correlations with objective tests and life experiences, cannot, of course, be decided on the basis of this study ; our expectation would be that inventory responses are more likely to prove valid indicators of the children's personality. Evidence from the literature (Rutter and Graham, 1966 ; Sarason, *et al.*, 1960 ; Wolff, 1967) suggests strongly that teacher's ratings, at least as far as neuroticism is concerned, do not correlate positively, and may correlate negatively with valid external criteria.

IV.—REFERENCES.

EYSENCK, H. J. (1960). *The Structure of Human Personality.* London : Methuen.

EYESNCK, H. J., and EYSENCK, S. B. G. (1968). *Description and Measurement of Personality.* London : Routledge and Kegan Paul.

EYSENCK, S. B. G. (1965). *The Junior Eysenck Personality Inventory.* London : University of London Press.

RUTTER, M., and GRAHAM, P. (1966). Psychiatric disorder in 10 and 11-year-old children. *Proc. Roy. Soc. Med.*, 59, 382-387.

SARASON, S. B., *et al.* (1960). *Anxiety in Elementary School Children.* New York : Wiley.

WOLFF, S. (1967). Behavioural characteristics of primary school children referred to a psychiatric department. *Brit. J. Psychiat.*, 113, 885-893.

(*Manuscript received 15th September, 1967*)

A VALIDATION OF THE HARRIS-GOODENOUGH TEST

BY D. J. W. STRÜMPFER
(*University of Port Elizabeth*)

AND C. J. P. MIENIE
(*Transvaal Education Department, Pietersburg, South Africa*)

SUMMARY. Scores on the Harris-Goodenough point scales were studied in a group of seventy-nine children (mean age 11·44 years). Various kinds of reliability proved to be acceptably high. Criterion measures were an individual and a group intelligence test, scholastic achievement tests, a paper-and-pencil dexterity test, the HSPQ, examination marks, rankings by teachers, and father's occupation. Validity coefficients ranged from low to moderate. The scales were poorer predictors of scholastic performance than conventional intelligence tests. Scores were not related more to non-verbal than to verbal performance. The only HSPQ scale which showed significant correlations was one indicative of a phlegmatic temperament.

I.—INTRODUCTION.

Since human figures are among the most popular subjects of spontaneous drawing and since children usually comply willingly with a request to 'draw a man,' the DAM is a convenient means for establishing rapport. Compared to the amount of time required for administration and scoring, it also provides much useful information.

The Harris-Goodenough Test of Psychological Maturity (Harris, 1963) is a revision and extension of Goodenough's (1926) original test. Restandardisation on contemporary and more representative groups was the first aim. An analogous point scale for female figures was also developed as an alternate form to the man scale. The age range was increased by two years at the top, so that the new scale can be used for children between 3 and 15. Scores are expressed as deviation I.Qs. Product scales, comparable to those of Burt (1921), were also developed; by means of these the quality of a drawing can be evaluated globally.

The present study concerned the reliability and validity of the point scales, i.e., those based upon drawing items of detail, proportion and perspective.

II.—METHOD.

Sample.

The sample consisted of seventy-nine white children (thirty-seven boys, forty-two girls) in standard four (sixth year) of an Afrikaans medium primary school in a large rural town. Their mean age was 11·44 years. They were predominantly from middle and upper-lower class backgrounds, but intellectually the sample seems to have been representative. On an individual intelligence scale (NSAIS) the overall mean score of this sample was 100·78 with a standard deviation of 14·97. The mean quotient from the verbal sub-scale was 99·29 (SD=15·70), and from the non-verbal sub-scale 102·51 (SD=12·82). Little reliance can be placed on this difference since it was not repeated on the group intelligence test sub-scales (NSAGT).

Criterion Variables.

The following criterion data were obtained:

New South African Individual Scale (NSAIS; National Bureau, 1964). Verbal, Non-Verbal and Total Scores were used, as well as scores for the subtests for the purpose of determing construct validity. The verbal subtests include Vocabulary, Comprehension, Reasoning, (Arithmetic) Problems, and Memory (retelling of paragraph). The non-verbal subtests include Pattern Completion, Block Design, (picture) Absurdities, and a Form Board.

New South African Group Test (NSAGT; National Bureau, no date). Verbal, Non-verbal and Total scores were obtained. The verbal subtests include classification of word pairs, reasoning and analogies. The non-verbal subtests are figure analogies, number series and pattern completion.

Scholastic Achievment Tests (Education Bureau, 1962). The tests of Afrikaans (first language), English (second language) and Arithmetic were used. The two language tests both contain subtests on vocabulary, comprehension, language usage and spelling. The arithmetic test covers the four main operations but does not include problems.

Junior-Senior High School Personality Questionnaire (HSPQ; Cattell and Beloff, 1962). An Afrikaans translation was used. Thirteen factorial scores were obtained (leaving out Factor B).

Dexterity Test (King, 1948). This is a paper-and-pencil test consisting of three sub-tests: following a maze; marking one-centimetre squares; and making dots in three-millimetre triangles. The tests are highly speeded, one minute being allowed per sub-test. The correlations between the sub-tests were significant but not high enough to have warranted their combination into a single score (·36, ·47 and ·51).

Examination Marks. Mean percentages obtained from seven subjects at the end of the previous year (four months prior to this testing) were used.

Ranking by Teachers. Six teachers who each knew the children well, first ranked them independently according to his or her evaluation of the child's scholastic ability. A final rank order was made after careful comparisons and discussion among the six teachers.

Father's Occupation. The occupations of the fathers were classified as: professional, administrative, clerical, farming, skilled, semi-skilled and unskilled labour.

III.—RESULTS AND DISCUSSION.

Reliability.

Odd-even reliabilities were determined for the Harris-Goodenough scales, although such a split does not result in strictly comparable halves. For the Man scale the Spearman-Brown corrected coefficient was ·88, and for the Woman scale ·92 (·78 and ·85 before correction).

The correlation between the Man and Woman scales was ·78 for the raw scores, and ·80 for the standard scores. This indicates that the Woman scale cannot be considered as fully equivalent to the Man scale. More reliable results can, however, be obtained by combining the two scales. Then the corrected internal reliability coefficients were ·87 for the combined raw scores and ·94 for the combined standard scores.

Sixty-nine children were retested four months after the initial testing. The test-retest reliability coefficients were ·73 for both the Man and Woman scales, and ·80 for the Total score.

The drawings were also scored independently by a second scorer; an inter-scorer reliability coefficient of ·91 was obtained for the Man scale, of ·92 for the Woman scale, and of ·94 for the Total score. These inter-scorer reliabilities are slightly higher than those reported by Auricchio (1966) for drawings by second-graders.

TABLE 1

CORRELATIONS BETWEEN HARRIS-GOODENOUGH QUOTIENTS AND CRITERIA OF INTELLECTUAL FUNCTIONING.

Criterion Variables	Harris-Goodenough Quotients		
	Man	Woman	Total
NSAIS:			
Verbal	·50	·37*c*	·45
Non-verbal	·61	·50	·58
Total	·58	·45	·54
NSAIS Subtests:			
Vocabulary	·47	·34*c*	·42
Comprehension	·44	·32*c*	·40
Reasoning	·45	·35*c*	·42
Problems	·38	·22*a*	·31*c*
Memory	·34*c*	·30*c*	·34*c*
Patterns	·35*c*	·25*b*	·31*c*
Blocks	·45	·44	·47
Absurdities	·51	·44	·49
Form Board	·32*b*	·21*a*	·27*b*
NSAGT:			
Verbal	·56	·54	·58
Non-Verbal	·43	·38	·43
Total	·54	·50	·55
Scholastic Tests:			
Afrikaans	·47	·43	·48
English	·57	·47	·55
Arithmetic	·24*b*	·25*b*	·26*b*
Examination Marks	·51	·41	48
Ranking by Teachers	·42	·43	·46

NOTE.—All coefficients significant beyond ·001 level, except when indicated differently.
a Not significant.
b Significant beyond ·05 level.
c Sign ificantbeyond ·01 level.

Validity.

No significant sex differences were found.

Correlations with the classification of the father's occupation were significant (·05 level) for the Man I.Q. (·27) and for the Total I.Q. (·22), with higher scores being associated with higher occupational background.

Table 1 shows the correlations between the Harris-Goodenough quotients and the various criteria of intellectual functioning. Comparing the correlations of the Harris-Goodenough Scales with verbal and non-verbal sub-scores, the NSAIS and NSAGT produce contradictory patterns. But the results in Table 2 indicate that these tests have rather low inter-correlations. Even with the NSAIS no consistent pattern of correlations emerged. Vocabulary, Comprehension, Verbal Reasoning, Block Design and Picture Absurdities showed the highest correlations, but this grouping cuts across the classification of verbal *vs.* non-verbal. This supports the conclusion of Harris (1962) that " the Draw-a-Man Test is not more allied with performance than with verbal abilities " (p. 99).

In general, the relatively low correlations between the Harris-Goodenough quotients and scores on the other intelligence tests are conspicuous. Even for the highest correlation only 37 per cent. of the variance could be explained in terms of a common factor. This picture also agrees with findings on the DAM summarised by Harris (1963). The present findings are especially similar to those obtained for the WISC, with which the NSAIS shows some similarity.

The relatively low correlations between the Harris-Goodenough scores and the more conventional intelligence tests are, of course, not really surprising. There is virtually no similarity between the two types of test, a point also made by Buck (1948), in connection with the quantitative scores on the House-Tree-Person Test. It is in part for this reason that Harris (1963) considered it advisable to substitute the concepts of intellectual or conceptual maturity for intelligence in a discussion of what is measured by the drawing test. Buck was of the opinion that the greatest value of quantitative drawing scores is to be found in the differential diagnostic information obtainable from a comparison of intelligence scores derived from these differing approaches.

The Harris-Goodenough quotients showed highly significant correlations with examination marks and teachers' ratings (Table 1). When the nature of the test task, as well as the amount of time required for testing, are taken into consideration, this is quite encouraging. These correlations are, however, well below those obtained between the NSAIS or the NSAGT and these criteria (Table 2), indicating once more that these tests are considerably better predictors of scholastic achievement.

The Dexterity Test was included to determine the importance of manual dexterity, as measured by a paper-and-pencil test, in performance on the Harris-Goodenough. Only the Maze test correlated significantly (·05 level) with the three quotients. The coefficients were ·26, ·25 and ·27. Eye-hand co-ordination and control of continuous voluntary movement are used in this test but it is clear that these are of minor importance in those aspects of drawing considered by the Harris-Goodenough scoring system.

Only one of the HSPQ scores, that for Factor D, showed significant (·05 level) correlations with the Harris-Goodenough quotients. The coefficients were −·28, −·25 and −·28. Cattell and Beloff (1962) describe this as a factor of " Phlegmatic Temperament *vs.* Excitability." A restrained, deliberate, constant temperament, therefore, appears to be slightly advantageous to performance on the Harris-Goodenough. On the other hand, it is clear that excitement, impatience, hyperactivity, distractability and impulsiveness could detract from a child's performance. It is, however, doubtful whether the 6 or 8 per cent. of the drawing variance which can be explained in terms of this variable is of any clinical significance.

With only one exception, the Arithmetic test, all the validity coefficients were higher for the Man scale than for the Woman scale. This confirms Harris's surmise that : " Possibly the Point scale method of scoring the woman figure is not quite as revealing of changes in psychological maturity as is the man figure " (1963, p. 119).

TABLE 2

INTERCORRELATIONS OF CRITERIA OF INTELLECTUAL FUNCTIONING.

Score	NSAIS		NSAGT			Scholastic Tests			Exam. Marks	Ranking
	N–V	T	V	N–V	T	Afr.	Eng.	Arith.		
NSAIS:										
Verbal	·71	·95	79	·61	·77	·77	·81	·53	·70	·69
Non-verbal		·89	·72	·64	77	·60	·66	·47	·68	·62
Total			·82	·67	·81	·75	·81	·55	·74	·72
NSAGT:										
Verbal				·66	·90	·85	·78	·58	·71	·75
Non-verbal					·91	·53	·55	·52	·64	·55
Total						·77	·73	·60	·75	·71
Scholastic Tests:										
Afrikaans							·85	·71	·71	·84
English								·63	·75	·84
Arithmetic									·67	·81
Examination Marks										·71

NOTE.—All coefficients significant beyond ·001 level.

In conclusion, the validity coefficients which were obtained do not seem to be contrary to the assumption that the Harris-Goodenough point scales provide measures of a child's ability to form abstract concepts. The relatively low validity coefficients make it a poorer predictor of scholastic performance than conventional measures of intelligence. Considering the ease of administration, the Harris-Goodenough, nevertheless, seems to justify a place for itself as a supplementary tool for the clinical or the educational psychologist, especially in situations where administration of another test is impractical.

ACKNOWLEDGEMENTS.—This study was partially supported by a grant from the Research Committee of the University of Port Elizabeth. We wish to thank Mr. G. K. Huysamen for his assistance with the scoring of drawings.

IV.—REFERENCES.

AURICCHIO, ELIZABETH W. (1966). Comparison of several methods of scoring draw-a-person tests. *Percept. Mot. Skills*, 23, 1124.

BUCK, J. N. (1948). The H-T-P technique; a qualitative and quantitative scoring manual. *J. Clin. Psychol.*, Monogr. Suppl., No. 5.

BURT, C. (1921). *Mental and Scholastic Tests.* London: King.

CATTELL, R. B., and BELOFF, H. (1962). *Handbook for the Jr.–Sr. High School Personality Questionnaire.* Champaign, Ill.: Inst. Pers. Ability Testing.

EDUCATION BUREAU (1962). *General Manual for Teachers for the Interpretation and the Use of the Results of the Scholastic Tests of the Education Bureau.* Pretoria: Transvaal Educ. Dep.

GOODENOUGH, FLORENCE L. (1926). *Measurement of Intelligence by Drawings.* New York: Harcourt, Brace and World.

HARRIS, D. B. (1963). *Children's Drawings as Measures of Intellectual Maturity.* New York: Harcourt, Brace and World.

KING, J. E. (1948). *Handbook for Use of the Factored Aptitude Series of Business and Industrial Tests.* Chicago: Industr. Psychol.

NATIONAL BUREAU OF EDUCATIONAL AND SOCIAL RESEARCH (1964). *Preliminary Manual: New South African Individual Scale.* Pretoria: Dep. Educ. Arts and Science, 1964.

NATIONAL BUREAU OF EDUCATIONAL AND SOCIAL RESEARCH (no date). *Manual for the New South African Group Test Intermediate Series, Form G.* Pretoria: Dep. Educ. Arts and Science.

(*Manuscript received 13th September, 1967*)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

THE EFFECT OF TRAINING IN STUDY SKILLS ON THE ABILITY TO RECALL AND COMPREHEND WRITTEN MATERIAL

By BLANCA AZPURUA-ARRILLAGA

(*Abstract of Thesis presented for the degree of M.Ed. in the University of Manchester*, 1966)

The aim of the research was to investigate the effect of training in study skills on the ability of secondary modern school pupils to comprehend and recall written material. Three groups of thirty-eight third-form pupils, aged 13 and 14 years, who were comparable in socio-economic level, scholastic ability, reading ability (Iowa Silent Reading Test), and academic interests, were matched for intelligence (Raven's Progressive Matrices). One group was given instruction in study skills, a second group instruction together with practice under supervision, and the third group acted as a control. The course of instruction lasted seven weeks with two periods per week. The groups were tested on their study of topics outside the school curriculum to avoid differences in training and practice effects. One topic was in the sciences (Pavlov and the Conditioned Reflex) and one in the Arts (A Brief History of Impressionism). The tests were given after studying each topic under controlled conditions, immediately and again three weeks later.

It was hypothesized that :

(1) Children of 13 and 14 years of age benefit from a course in study skills.

(2) The groups taught Study Skills have better comprehension and recall more of what they study than the groups not taught these skills.

(3) The actively trained group will comprehend and retain more than the group who have received instruction only.

(4) The experimental groups will retain proportionately more of what they have learnt than the control group.

An analysis of variance was performed to separate the effects of groups (two experimental and one control), levels of intelligence (two levels), topics (two), and occasions of testing (two), together with the six first-order interactions, four second-order interactions, and one third-order interaction between these effects. Hypotheses 1 and 2 were confirmed, the groups receiving instruction performing significantly better than the control group. Hypothesis 3 was not confirmed, the group receiving only instruction performing slightly (and significantly) better than the group receiving instruction plus practice. Hypothesis 4 was also not confirmed, for although significantly more was retained by the groups receiving instruction, they forgot proportionately more of what they had learned than the control group. It is recommended that courses on " How to Study " form part of the curriculum of secondary modern schools, and that more research be made to determine the need for such courses in grammar schools.

(*Manuscript received 21st June*, 1967)

MARRIED WOMEN AND YOUNG WOMEN STUDENTS AT A DAY COLLEGE OF EDUCATION AND IN THEIR FIRST YEAR AS TEACHERS

By DUNCAN CASE

(*Abstract of a thesis submitted for the degree of M.Ed. at the University of Leicester,* 1967)

I.—Aims.

This research was designed to examine and compare the attitudes and interests of mature married women and young single women undergoing a three-year course of training at a day teachers' college and in their first teaching posts.

II.—Samples.

Cross-sectional samples were used as follows: 220 female students attending the day-college of education (130 married, 90 single); 247 probationary teachers representing 70·6 per cent. of the sample to which questionnaires were sent; 293 teachers and non-teachers (aged between 20 and 60 years) from which two groups of sixty-four were matched by age and background.

III.—Measures.

(1) A questionnaire was devised to obtain biographical data, which included marital status, size of family, and domestic and professional commitments.

(2) An instrument (ENVA Test) consisting of thirty items was devised to test 'cultural-environmental awareness.'

(3) Richmond's Two-cultures Test No. 2.

(4) Four attitude scales were constructed to measure level of aspiration, educational progressivism, career satisfaction and job worry. The data thus obtained were submitted to a Guttman scalogram analysis.

(5) A questionnaire was devised to obtain information relating to the previous teaching experience, training, placement and sources of guidance and information of the teachers in their first posts.

IV.—Method.

A comparison was made between matched groups of teachers and non-teachers to test the hypothesis that the decision to teach would be related to a higher degree of cultural-environmental awareness. Comparisons between married women and single women were designed to investigate the possibility of role conflict for the married women at college and in their first-year of teaching. Finally, comparisons were made between mature married women, young married women and single women in their attitudes to teaching and in their leisure pursuits.

V.—Discussion of results.

It was found that cultural-environmental awareness did increase with age, but this test did not discriminate between the groups of teachers and non-teachers. These ENVA scores for the married women teachers in the sample correlated highly (average $r = \cdot 83$) with the total score on the Richmond Two-culture Test. The scores of these married women on Richmond's test were closer to the norms (Richmond, 1963) of final-year university students than to those of training college women. These high scores may be taken as showing an alertness which is in keeping with the ability of these women to occupy and manipulate successfully their multiple roles and confirms in a general way the findings of Soloman (1965). These married women appeared to cope with their multiple roles by using a system of role priorities.

In their attitudes towards teaching, the married women were seen to be only slightly less permissive, but not less committed, than the single women. High levels of professional satisfaction were expressed by both married and single women together with a deep concern, if not worry, over the job of teaching. This result seems to confirm the evidence of Rudd and Wiseman (1962).

During their probationary year, the mature married women appeared to be more stable and somewhat more desirous of improving themselves professionally than the young single and young married women in the sample. All these probationary teachers tended to rely on their head teachers for guidance, but the younger women tended to use only the teachers' popular magazines as their sources of information. The range of indicated leisure activities was wide and impressive, but the mature married women showed a greater inclination towards cultural pursuits than the younger married women, who were more concerned with home-making.

In general these results are taken as underlining the determination and the organisational qualities of the mature married women in this sample.

VI.—REFERENCES.

RICHMOND, W. K. (1963). *Cultural and General Education*. London: Methuen.

RUDD, W. G. A., and WISEMAN, S. (1962). Sources of dissatisfaction among a group of teachers. *Brit. J. Educ. Psychol.*, 32, 275-291.

SOLOMAN, EDITH (1965). *Personality Factors and Attitudes of Mature Training College Students*. Unpublished M.Ed. thesis, Manchester.

(*Manuscript received 25th August*, 1967)

A COMPARATIVE STUDY OF THE USE OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN AND THE STANFORD-BINET INTELLIGENCE SCALE (FORM L–M) WITH EIGHT-YEAR-OLD CHILDREN

BY MICHAEL BRITTAIN

(*Abstract of a dissertation presented for the degree of B.Tech. of Brunel University*)

I.—SAMPLE.

Forty-one children (twenty-two boys, nineteen girls) selected at random from the 8-year-old population (mean age 8 years 6 months) of one South-coast urban local education authority were tested with the WISC (not including the alternative tests) and form L–M of the S–B.

II.—MEANS AND STANDARD DEVIATIONS.

The group means and standard deviations were, for WISC (FS) m=108·6, s=16·8; and S–B m=104·4, s=17·5. This difference is not statistically significant, but its direction is contrary to that consistently found in comparisons involving Form L of the S–B (Littell, 1960; Jones, 1962).

III.—SEX DIFFERENCES.

Significant differences between means (i.e., p ·05) were found for S–B, the main WISC scales and for five of the WISC tests; viz., Arithmetic, Vocabulary, Picture Arrangement, Block Design, and Object Assembly. In each case the mean score for boys was higher than that for girls.

IV.—PRACTICE EFFECTS.

No significant (p ·05) practice effects were found in this study. The findings of Kureth (1952) and Holland (1953) using S–B Form L—that there are no significant effects upon WISC scores when S–B is administered first—are supported by the results of this study, using S–B Form L–M. The rise in mean S–B score when the

WISC is administered first, although not significant, was marked (a difference of six points between means) and this aspect may well be worth further study.

V.—VALIDITY AND RELIABILITY.

Correlations between S–B and WISC were WISC (FS) 0·89, WISC (V) 0·81, and WISC (P) 0·82. Reliability co-efficients calculated according to Kuder-Richardson Formula 8 ranged from 0·75 to 0·90 for individual WISC tests, and for the main scales were 0·94 (S–B), 0·93 (WISC FS), 0·92 (WISC V) and 0·91 (WISC P).

VI.—ITEM DIFFICULTY.

The order of items in both tests was shown to be inappropriate for this sample. The discrepancies between the published item order and the actual order of difficulty could be accommodated in the case of WISC by allowing four consecutive errors on the Comprehension test, five consecutive errors on the Similarities test and eight consecutive errors on the Vocabulary test before ceasing testing on the respective tests. It should be noted that the only correct answer to Picture Completion test item 13 is ' Nothing ' and this has been acknowledged by the Psychological Corporation. Discrepancies in S–B are not sufficiently great to require any alteration to the published instructions for administration.

VII.—REFERENCES.

HOLLAND, G. A. (1953). A comparison of the WISC and Stanford-Binet I.Qs. of normal children. *J. Consult. Psychol.*, 17, 147-152.

KUDER, G. F., and RICHARDSON, M. W. (1939). The theory of estimation of test reliability. *Psychometrika*, 2, 151-160.

KURETH, GENEVIEVE, MUHR, JEAN P., and WEISBERGER, C. A. (1952). Some data on the validity of the Wechsler Intelligence Scale for Children. *Child Developm.*, 23, 281-287.

LITTELL, WILLIAM M. (1960). The Wechsler Intelligence Scale for Children—Review of a decade of research. *Psychol. Bull.*, 57, 132-156.

PSYCHOLOGICAL CORPORATION (1967). *Personal communication.*

(*Manuscript received 22nd November*, 1967)

BOOK REVIEWS

CONNELL, W. F., DEBUS, R. L., and NIBLETT, W. R. (1967). *Readings in the Foundations of Education.* London: Routledge and Kegan Paul, pp.xvi+398, 42s.

In one of the readings from this collection Arnold Toynbee writes about the problem of coping with the cultural accumulation that is stuffed into the limited capacity of formal education. "How is a limited human mind to cope with a cultural heritage that is perpetually increasing in bulk? . . There will be a temptation to try to facilitate the acquisition of the growing heritage by simplifying its content at the cost of impoverishing it." The compilation of a set of readings in the foundations of Education is one device for attempting to convey to students some sense of the range of important educational problems that have been studied while preserving some of the original flavour of the scholars who have studied them.

The readings are organized in eight sections: (1) the aims of education; (2) the scope of education; (3) education and society; (4) special problems in contemporary education; (5) intellectual development; (6) social-emotional and personality development; (7) learning and teaching; and (8) planning and organizing the curriculum. They are intended to present basic and significant statements of leading issues and fundamental ideas and to have the English-speaking world as their realm of reference. Readers will find mostly rather good extracts from a wide range of educational writers—Bernstein, Bruner, Cronbach, Dewey, Guilford, Hoggart, Kilpatrick, Mannheim, Newsom, Peel, Pius XI, Rousseau, Shapovalenko, Snow, Vernon, Whitehead, Wiseman, and Wordsworth, to mention a non-random few—and these would certainly give students some introduction to philosophical, sociological, psychological, and organizational approaches to important educational problems. Many of the extracts are short but others are substantial enough to avoid the shattered-pottery effect that such volumes may generate. The brief commentaries that introduce sections and individual extracts have an admirable simplicity and lucidity, unobtrusively sustaining the flow of the book. But does this compendium simplify *without* impoverishing?

On the negative side one might ask whether a little more showing might have been given to the political and economic foundations of education, to more radical philosophical and sociological challenges to traditional thinking, or to the non-communist materialism and secularism which is the seriously practised, if not significantly stated, philosophy of so many thinking as well as unthinking people. On the positive side, the book must be one of the best of its kind—intelligent and sensitive in plan and execution, a handy reference and source volume for lecturers (and examiners) in Education, a good tutorial-course reader for students, and good value for money for whoever buys it.

H. S. N. MCFARLAND.

FLINT, B. M. (1966). *The Child and the Institution.* London: University of London Press, Ltd., pp. xii+177, 30s.

The work described in this book started in 1957 as a research project and developed into an 'experiment in human living.' It sought to test the extent to which the effects of extreme 'maternal deprivation' were remediable. The Institute of Child Study of the University of Toronto was 'academically interested' in the matter and the Catholic Children's Aid Society had practical problems in connection with it. Co-operation between the two bodies led to the project described. The Neill McNeil Home, where the work took place, is characterised by Professor Flint as a 'typical institution.' However, it would seem initially to have been one of the worst of its kind. It catered for infants and pre-school children and was overcrowded, understaffed and unsuitably housed. Contact between children and staff was superficial, arrangements and discipline were rigid and even 'severely damaging,' while poor mental and physical health were obvious among the children, many of whom were reduced to near-psychosis.

Professor Flint and her admirable colleagues set out to alter all this and to change the Home to a 'treatment centre.' They generated sufficient enthusiasm among those practically responsible for a massive and expensive (but ultimately economic) programme to be carried through. Fundamental was the provision of a better staffing ratio, with more psychologically aware staff, but building operations had also to be undertaken to improve kitchens and wards and to provide a play room. The latter was particularly important in contributing to the observation and therapy of the children which is fully described.

One item which, in retrospect, seems to have made a crucial contribution to the treatment plan was, however, free. It consisted in the organisation of a body of 'volunteer mothers.' These at one time numbered eighty. Their help was not casually deployed. They underwent a degree of training, had clear instructions, and were expected to keep formal records. The basic idea was that each child would benefit from having someone who would be "interested in him alone and to whom he could feel he belonged."

A feature of the work, deriving from its research character, was the keeping of adequate records. The book gains much from this. The children were periodically vetted and observed by psychologists and child development experts. Professor Flint's Infant Security Scale and Pre-School Mental Health Assessment Scale were the main instruments used. But full and part-time staff, as well as volunteers, prepared special questionnaire records at intervals. Guidance and management of the children were related to these and when a child seemed ready for fostering or adoption, careful preparation and counselling of the prospective home was given. Some detailed and very informative case histories show the way in which all this material was co-ordinated.

The book is interestingly and precisely written and should be of great value to those working in the field of child care. One looks forward to the promised longitudinal studies following up the children and studying their later history. Will the 'remarkable recovery' be maintained?

Olive C. Sampson.

Hargreaves, D. H. (1966). *Social Relations in a Secondary School.* London: Routledge and Kegan Paul, pp. xii+226, 32s.

It could not, and should not, be hoped that in two hundred odd pages it would be possible to present a comprehensive review of the social relations of a secondary modern school of 450 boys and twenty-six teachers. This is the report of a Government financed empirical study, in which from the outset the author-investigator adopted the more reasonable target of concentrating upon about a hundred boys in the fourth year, and studying behaviour and opinions indicative of attitudes towards the educative process. These he relates to a major structural feature, the streams in which the boys are organized. The data gathered from questionnaires and sociometric inventories are supplemented from the author's observations gleaned during a year's participant observation in the school: the problems of the various vantage points he tried are discussed frankly in an appendix.

The school studied appears to orient education towards examinations to a greater extent than is usually thought desirable nowadays. The boys are streamed by the results of examinations held every term. Although Hargreaves denies that this is primarily a study of streaming, his main conclusion is that this very grouping and segregation of boys in streams, and the academic criteria which decide allocation, create two sub-cultures within the school, polarised in relation to academic values—ultimately important values of society. The negative pole, located in the lower streams, he calls a 'delinquescent' subculture, arguing that this is a breeding-ground for values which in a suitable milieu would result in actual delinquency: indeed, over half the D stream admitted to current petty thieving.

One could carp at certain methodological points, such as the uncritical way in which conclusions are drawn from sociometric questions, and the failure to discuss the validity of the author's selection of comments and interpretations to support data under review: a merit of the clarity of presentation is that such problems are readily apparent. Clearly, with so important a conclusion before us, it is necessary to be sure that the case is well and truly made. There must be misgivings that the submitted polarization of attitudes is as much a product of the initial hypothesis and the observations and questionnaire items it engendered than of the boys' own value systems. (Even the decision to omit the fifth stream suggests a refusal to find a solution that fits a normal curve of distribution.) The author himself acknowledges the possibility of over-polarization, and attempts to support his case with somewhat slender data from the contemporary second year classes, suggesting the dynamics of a polarization. But this, as he recognizes, is an inadequate substitute for studies of change in the same individuals. His statistical case rests largely on grouped data which encourage him to overstate his position when they fit his hypothesis, and, when they don't, give off a resounding clang.

However much one may feel privately that the dangers of a school organization such as he describes are real, however much one might like to see similarly identified the counterpart of the 'delinquescent' subculture among the non-conforming intelligentsia of some selective secondary schools, scepticism is desirable. But it should be the sort of scepticism that asks for more, not less, work of this kind—and work which, while gratefully acknowledging imaginative pioneering such as this, is on a scale sufficient to eliminate the concatenations inevitable in a one-man band.

THELMA VENESS.

HOLLOWAY, G. E. T. (1967). *An Introduction to the Child's Conception of Geometry* (pp. 71, 6s.) and *An Introduction to the Child's Conception of Space* (pp. 77, 6s.). London: Routledge and Kegan Paul.

These handbooks provide a simplified account, mainly by direct quotation, of 'The Child's Conception of Geometry' by Piaget, Inhelder and Szeminska, and 'The Child's Conception of Space,' by Piaget and Inhelder. Each handbook has fewer than 80 pages; they enable a reader to see quickly Piaget's analysis of geometrical and spatial concepts, and to appreciate the stages by which these concepts develop in children.

All the experiments Piaget uses are described, but occasionally the details are condensed to such an extent that the reader is in doubt about what is required of the children tested. For example, more details should have been given of the experiments on knots and on similarity of triangles in the space handbook. A part of the discussion on measurement of lengths (geometry handbook, p. 23) mentions three objects named M, A, B; this part can be understood only by reference to the original book, p. 122.

On p. 56 of the geometry handbook, it is implied, mistakenly, that stage 4 of the measurement of areas is reached between 8 and 9 years of age. Elsewhere in the handbooks, stage 4 is used, following Piaget, as another way of referring to the stage of formal operations.

In books written as summaries of existing books, it is not easy to judge the appropriate degree of explanation to be given to technical terms. Phrases occur in the space handbook, such as "assimilated to the schema of" (p. 8) and "concrete operations" (p. 66) and in the geometry handbook such as "multiplicative logical operations" (p. 71). Readers who study the handbooks before reading Piaget's works will require an explanation of such terms. Only one such term is described—that of an operation (space handbook p. 7). This description does not state that an operation is an action in thought, an internalised action. Without enlarging the handbooks an explanation of some technical terms could well have been given as in the geometry handbook there are over five empty pages, and in the other over three empty pages, which serve merely to separate their sections.

While these handbooks will serve as a very useful introduction to students and teachers to the original books on geometry and space, it seems a pity that the opportunity was not taken to discuss the wider implications of concept analysis and guided activity in schools.

J. REYNOLDS.

HOLT, K. S., and REYNELL, J. K. (1967). *Assessment of Cerebral Palsy.* Vol. 2., Lloyd-Luke (Medical Books), Ltd., pp. viii + 187, 40s.

In the previous volume, Dr. Holt wrote on the assessment of motor handicap in cerebral palsied children. In the present book he and Dr. Reynell deal with four further areas of possible impairment, namely, speech and language, hearing, vision and psychological function. The discussion of each of these topics follows a similar pattern. The authors first outline the functions affected by these types of disability, and, in varying detail, describe the mechanisms underlying them. In addition, they provide summaries of normal and defective development. They then consider diagnostic procedures, and conclude by making suggestions for the effective communication of the diagnostic findings.

Although these topics are considered in the context of cerebral palsy, the book, in fact, provides a unique summary of these disabilities themselves. Considering the length of the book (187 pages) the authors have collected together a wealth of information which, up until now, one could only find scattered about in a wide variety of detailed books. The selection and presentation of this information is an achievement in itself. Perhaps even more important is the view of the handicapped child which the authors offer. They show how the consequences of disability cut across the conventional distinctions between medicine, psychology and other related fields. They stress that those serving the handicapped child need to bear this in mind, particularly in attempting to diagnose multiply handicapped children. It is often hard to discover which of a child's stated disabilities presents the greatest handicap to him. It is not surprising, therefore, to find the authors pleading that diagnostic findings should reach those who are actually responsible for the child, in a formulation which is of practical value to them.

The sections on psychological assessment echo the present dissatisfaction with the use of global intelligence test scores. The authors point to the variety of causes which may underly failure even on a single test item, and describe means of assessing some of these types of specific intellectual impairment. Although general reference to the diagnosis of educational failure is made throughout the book, one would have appreciated more detailed information. Perhaps this could be the subject of a further volume, along with an extension of the chapter on social and emotional development.

This book presents an outstanding introduction to the assessment of handicapped children. Some of its statements appear too categorical, but the authors provide useful references for further reading, and amplify their points by many case histories.

The book is amply illustrated and is almost free of misprints. The layout is clear but unoriginal The references are tiresomely arranged in non-alphabetical order at the end of each chapter.

K. WEDELL.

MAIER, H. W. (1965). *Three Theories of Child Development.* New York: Harper and Row, pp. 314, 30s. (*Harper International Students' Reprint.*)

This book is sub-titled *The Contributions of Erik H. Erikson, Jean Piaget, and Robert R .Sears and their Applications.* Its purpose is threefold : first, to set out the views of these three writers in comparable terms ; second, to bring these views together and, as far as possible, to establish a synthesis of all three ; and, third, to examine the aims and methods of members of ' the helping professions,' i.e., of therapists of various kinds, and to show how these can be furthered by such a comprehensive theoretical reference frame. These aims, laudable as they are, were perhaps too ambitious to be entirely successful. The main part of the book is devoted to the exposition of the three theories. These accounts are rather brief, and although accurate in detail, they offer somewhat distorted perspectives of the authors they seek to represent. This is because, Erikson excepted, the writers are not primarily concerned with the ' helping ' process, while Maier undoubtedly is. Nor is comprehension made easy by the transatlantic style of writing (by Kant out of Johnson). Maier's book will not be very useful to the student who is looking for a compact introduction to Piaget. The treatment of Erikson and Sears is rather more successful, but these authors also are more readable in the original.

Perhaps the most thought-provoking section of the book is Maier's own analysis of the 'helping' process, which contains many useful insights for clinical students. At any rate, this reviewer found it more stimulating than either the synthesis or the attempt to show the application of theory. Maier's volume has much to interest a thoughtful reader, but it cannot be recommended as a standard textbook.

E. A. LUNZER.

PAEDAGOGICA EUROPAEA: Vols. 1965, 1966. Edinburgh and London: W. and R. Chambers, Ltd., pp. xii+284, xiv+339. no price.

This international yearbook sets out to perform the useful function of providing a European forum for the exchange of ideas and research results in education, and possibly to foster the co-ordination of research projects. With the major articles which form the greater part of its contents, it includes information about centres for research in different countries, statements of international projects already in operation, and reports of resolutions agreed at conferences of the European Ministers of Education. (This part, though well intentioned, may prove less satisfying than expected, since the information is necessarily brief and general, and official resolutions seldom rejoice the reader by their precision, density of information or novelty of ideas.) The interest of the book is mainly for educationists, especially, perhaps, those concerned with comparative education, but the contributions, while showing how importantly education is influenced by factors other than psychological theory and research, do include topics of interest to psychologists, e.g., psycholinguistic approaches to the study of development (Bruce, 1965), teaching of thought operations (Kohnstamm, 1965), prognostic value of selection techniques for the German Gymnasium (Knoche, 1965), sex differences in development of moral concepts (Lloyd, 1966). (And indeed Kohnstamm's emphasis on assessing the child's ability to profit not simply from spontaneous or reinforced learning, but from educational explanation in the test situation might yet reconcile psychologists with teachers who coach for intelligence tests.) Many of the comments on research and its methods are of interdisciplinary interest; thus, Langeveld (1966), makes a valuable analysis of the discontinuity from which research has suffered and the fashion of disregarding earlier work which is not venerably antique or safely modern; his illustrations from the topics of maternal deprivation and linguistic development are effective. (His article in the 1965 volume also offers admirable general principles.)

As seems inevitable in a yearbook, some articles are of more value than others; but a reader's attempt to read selectively is not greatly helped by the summaries in two languages; one knows the problems of compressing (the prize here is won by the three-quarter page summary of Moser and Layard's fifty-page-long analysis (1965) of statistical problems involved in producing the Robbins' Report and similar studies), but such summaries conceal whether the article contains anything of real value. One has to be able to read in English, French and German, to be sure of not missing something. (This is perhaps a device to develop an essential attribute of the European of the near future, just as the rebarbative title may have been the only way of avoiding the difficulties discussed by Landsheere (1965) apropos of producing an international dictionary of educational terms.)

In all, this publication does give a comprehensive and intelligent view of education's many aspects.

MARGARET B. SUTHERLAND.

SEGAL, S. S. (1967). *No Child is Ineducable.* Oxford: Pergamon Press pp. xii+332, 30s.

The sub-title of this book, *Special Education: Provision and Trends*, very adequately describes the contents. Although primarily meant for professional workers interested in the field of special education, the author has also written for intelligent parents. Segal is concerned with all forms of handicap in the school-age child. He recognises the value of an inter-disciplinary approach and summarises information from a very wide field. He demonstrates the multiple causation of backwardness, but stresses that all handicapped children require education. The author writes cogently, for example, of the need to bring the training of severely subnormal children under the influence of education.

The book is divided into three sections. The first part describes present-day provisions for the educationally and mentally subnormal within the framework of historical development. Three chapters are concerned with cognitive assessment and particularly with the 'inconstancy of the I.Q.' The author rightly stresses the dangers of attaching labels to children, but his discussion of cognitive assessment seems to lack balance. There are real dangers in un-skilled testing and naive interpretation, and for this reason it is a pity that little is said of the merits of cognitive assessment. Language development is briefly discussed but here again the presentation is superficial. The author refers to a wide selection of research and is concerned to recommend it to the practitioner, but does not attempt to describe the work in any detail. One minor irritation is the occasional mis-spelling of authors' names and the omission of some publication dates within the text.

The second part of the book, which makes up approximately a third of its length, is a guide to voluntary and professional organisations concerned with handicap and disorder. The main aims and functions of these societies are indicated and this section should be of considerable practical value to all those working in the field of special education.

The third and shortest section indicates some of the changes which are taking place in special education, and discusses the need for research and specialised training for teachers dealing with handicapped children.

In general, the book demonstrates the growing optimism of those working with slow learning children. The book might have been better divided into two volumes : (1) a guide to social and educational services and (2) an introduction to the assessment and education of handicapped children. The author has, perhaps, fallen into the trap of writing for too diverse an audience. Nevertheless, the book should make a useful addition to the library.

R. I. Brown.

Tallent, N. (1967). *Psychological Perspectives on the Person.* London : D. Van Nostrand Compay, Ltd., pp. 292, 24s.

The stated aim of this book is to enable its readers to achieve better insight into themselves and others—the main reason, according to its author, for students joining psychology courses.

As might be expected, therefore, the viewpoint is frankly clinical—" Freud has contributed more than any other scientist to the understanding of human personality "—but eclectic. Operational definition and objective measurement are heavily attacked and are shown to be inadequate to assess self-actualization and wisdom. This last attribute receives virtual deification and is referred to in italics ! " What a pity we don't know better how to define *it*, to identify better those who have *it*, or who particularly have the capacity for developing *it*, for cultivating *it*."

This quotation illustrates clearly the main weakness of the book. British psychology, generally, attempts to be scientific and is concerned with quantification and experimental design whereas the emphasis in this book lies in personal insight, on knowledge in the Platonic sense. On the other hand, this approach ensures that the book is always relevant to human, as distinct from rodent, behaviour, and so insightful are some of the examples of human behaviour that there is a real need for the warning, which the author gives, concerning identification with the various types of individual portrayed, the phenomenon well known to perusers of medical texts. Although interpretation of behaviour leans heavily on Freudian concepts in certain places, the chapter on controlling emotions, for example, being a discussion of psychoanalytic defence mechanisms, in other chapters quite different work is described. Thus, on the social side, much is made of the work of Leary, whose results, however, are not integrated with other views examined in the book. Thus, there is a tendency for the reader to be left at the end of the whole with a number of quite different, yet each apparently insightful and valid, vignettes of human motivation—the bitter fruits of eclectism ?

There can be little doubt that a psychologically naive reader will better understand some human behaviour (if clinical views are yielded any validity) so that the first aim of the book is achieved. On the other hand, it could be argued that a simple exposition of the various clinical doctrines discussed could have done the same. Yet this is hardly fair to the author whose easy flowing style, interlaced with humour, makes the book very pleasant to read.

PAUL KLINE.

TANSLEY, A. E. (1967). *Reading and Remedial Reading.* London: Routledge and Kegan Paul, pp. xii+176, 20s. net.

The material for this book was originally compiled for the benefit of the school staff of a residential school for educationally subnormal children to enable them, one assumes, to understand more fully the principles guiding their school and to cope more efficiently with the specific learning problems facing them. It is now offered to a much wider audience including those tackling the teaching of reading in normal schools. Some dangers immediately become apparent. What is suitable for a small, highly skilled group of teachers may not always have universal application. For instance, Mr. Tansley advocates the use of certain tests (e.g., the Bender Visual Motor Gestalt test and the Benton Visual Retention test) which have a 'c' level classification from the NFER and so may not be so readily available as he implies.

The book is divided into two parts, one dealing with the teaching of reading with particular emphasis on reading readiness and initial approaches and a second part which tackles the remedial teaching of reading. The separation of the two aspects is, however, far from complete as the author maintains at the very beginning that the learning processes of backward and bright children are the same: the difference lies in speed of learning. It is difficult to reconcile this contention with his other equally firmly expressed statement that "reading is a psychoneurological process and any structural or mechanical damage to or disturbance in the neural pathways will lead to specific symptoms and necessitate special types of remediation" (p. 134). If psychoneurological malfunctioning is stipulated as being a major source of reading difficulties, it does seem that the learning processes of the children so affected may differ in other aspects than speed.

Relying for theoretical support on the work of Hebb, Delecato and Lise Gellner, Mr. Tansley outlines an interesting diagnostic programme for poor readers, aimed at precise definition of the motor or perceptual modalities affected. Highly practical training programmes are listed and case studies quoted to prove their worth. Teachers or intending teachers of children who are finding difficulty in learning to read will find the book extremely helpful and stimulating, provided they appreciate the speculative nature of much of what is proposed and do not assume that all underachievement in reading has a neurological basis.

It is a pity that Mr. Tansley has relied so exclusively on empirical observation, however extensive, and has allowed his own convictions to feature so largely. There is plenty of material for discussion here.

A. F. LAING.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. The date of publication is 1967 unless otherwise stated. "n.p." means "no price stated."

Ablewhite, R. C. *The Slow Reader*. London: Heinemann, pp. vii+95, 15s.

Alexander, F., and Selesnick, S. *The History of Psychiatry*. London: Allen and Unwin, pp. xvi+472, 63s.

Apter, M. J. *Introduction to Psychology*. Bristol Teaching Programmes, pp. 625, 25s.

Aronoff, J. *Psychological Needs and Cultural Systems*. London: Van Nostrand, pp. 239, 24s.

Beard, R. M. *Research into Teaching Methods in Higher Education*. London: Society for Research into Higher Education, pp. 43, 12s. 6d. (8s. to members).

Beinum, H. Van. *The Morale of the Dublin Busmen*. London: Tavistock, pp. 94, n.p.

Belson, W. A. *The Impact of Television*. London: Crosby Lockwood, pp. 400, 50s.

Biggs, J. B. *Mathematics and the Conditions of Learning*. London: N.F.E.R., pp. xi+441, 55s.

Blackham, G. J. *The Deviant Child in the Classroom*. California: Wadsworth, pp. xii+208, $6.60.

Boothman, D. B. *Topical*. London: Longmans, pp. 105, 12s. 6d.

Boss, P. *Social Policy and the Young Delinquent*. London: Routledge, pp. 102, 15s.

Brearley, M., and Hitchfield, E. *A Teacher's Guide to Reading Piaget*. London: Routledge, pp. 171, paperback 8s. 6d.

Bresston M. *The American Male*. London: Allen and Unwin, pp. 252, 28s.

Cohen, J. *Psychological Time in Health and Disease*. Illinois: Thomas, pp. xv+101, $6.50.

Connell, W. F., Debus, R. L., and Niblett, W. R. *Readings in the Foundations of Education*. London: Routledge, pp. xvi+398, 42s.

Copel, S. L. *Psycho Diagnostic Study of Children and Adolescents*. Illinois: Thomas, pp. xii+201, $7,75.

Cowan, G. *Project Work in the Secondary School*. London: Longman, pp. 91, 10s. 6d.

Crothers, E., and Suppes, P. *Experiments in Second Language Learning*. New York: Academic Press, pp. ix+374, £5 8s.

Cryer, B. J., and Kitt, I. *The Progress Books Series*. London: Longmans, Set A 12s. 6d., Set B 15s., Set C 21s.

Dahrendorf, R. *Conflict after Class*. London: Longmans, pp. 33, 6s.

Davis, C. *Room to Grow*. London: U.L.P., pp. xxvi+214, 30s.

Dixon, J. *Houses and Housing*. London: Longmans, pp. 152, 12s. 6d.

Eggleston, S. J. *The Social Context of the School*. London: Routledge, pp. xii+116, 15s.

Fairweather, G. W. *Methods for Experimental Social Innovation*. New York: Wiley, pp. x+250, 60s.

Flint, B. M. *The Child and the Institution*. London: U.L.P. (1966), pp. xii+177, 30s.

Furneaux, W. D., and Gibson, H. B. *The New Junior Maudsley Inventory*. London: U.L.P. (1966), n.p.

Geleerd, E. R. *The Child Analyst at Work*. New York: International Universities Press; Bailey and Swinfen, pp. 310, 63s.

Goldstein, B. *Low Income Youth in Urban Areas*. New York: Holt, Rinehart, Winston, pp. vii+280, $3.95.

Gosden, P. H. *Educational Administration in England and Wales*. Leeds: University Press, pp. 55, 6s.

Gowan, J., *et al.* (eds.). *Creativity: its Educational Implications*. New York: Wiley, pp. xii+336, 64s.

Gueron, G., *et al.* *Éducation sans Frontières*. Paris: Presses Universitaires, pp. 187, F.9.

Hildum, D. (ed.). *Language and Thought*. London: Van Nostrand, pp. 201, 16s.

Holloway, G. *An Introduction to the Child's Concept of Geometry*. London: Routledge, pp. 71, 6s.

Holloway, G. *An Introduction to the Child's Conception of Space*. London: Routledge, pp. 77, 6s.

Holmberg, B. *Correspondence Education*. Malmö: Hermods-Nki, pp. 78, 17s.

Horst, P. *Psychological Measurement and Prediction*. London: Prentice-Hall (1966), pp. xii+455, 72s.

Jacobson, E. *Biology of Emotions*. Illinois: Thomas, pp. xvii+211, $7.50.

Jakobovits, L., and Miron, M. (eds.). *Readings in the Psychology of Language*. London: Prentice-Hall, pp. 636, 80s.

Jeffreys, M. V. C. *John Locke—Prophet of Common Sense*. London: Methuen, pp. 120, 12s .6d.

Johanesson, I. *Effects of Praise and Blame*. Stockholm: Almquist and Wirsell, pp. 82, n.p.

Jones, J. K. *Colour Story Reading—A Research Report* (together with set of Colour Story Reading material). London: Nelson, £2 10s. (complete set).

Kelsall, R. K. *Population*. London: Longmans, pp. 113, 10s. 6d.

Kidd, A. J., and Rivoire, J. L. *Perceptual Development in Children*. London: U.L.P., pp. 548, 100s.

Leijonhielm, C. *Colours, Forms and Art*. Stockholm: Almquist and Wirsell, pp. 184, n.p.

Leon, A. *Histoire de L'enseignement en France*. Paris: Presses Universitaires, n.p.

Lyall, A. *History Syllabuses and a World Perspective*. London: Longmans, pp. 171, 11s. 6d.

McLeish, J., and Higgs, G. *An Inquiry into the Musical Capacities of Educationally Sub-Normal Children*. Cambridge Instititute of Education, pp. 12, 2s. 6d.

Mialaret, G., and Pham, D. *Statistique à L'Usage des Éducateurs*. Paris: Presses Universitaires, pp. 264, F.14.

Mair, A. *Student Heath Services in Great Britain*. Oxford: Pergamon Press, pp. xii+236, 42s.

Meltzer, D. *The Psycho-Analytical Process*. London: Heinemann, pp. xvii+109, 20s.

Millenson, J. R. *Principles of Behavioral Analysis*. New York: Collier-Macmillan, pp. xx+487, 63s.

Montgomery, G. W. G. *Vocational Guidance for the Deaf*. Edinburgh: Livingstone, pp. 42, 7s. 6d.

Morgan, J. S. *Managing the Young Adult*. New York: American Management Association; Bailey and Swinfen, pp. 207, 68s.

Morris, B. *Understanding Children*. Wellington: Reed; Bailey and Swinfel, pp. 177, 16s.

Morrish, I. *Disciplines of Education*. London: Allen and Unwin, pp. 336, 40s.

Moustakas, C.: *Creativity and Conformity*. London: Van Nostrand, pp. 142, 16s.

Nagera, H. *Vincent Van Gogh: a Psychological Study*. London: Allen and Unwin, pp. 182, 36s.

Passow, A., Goldberg, M., and Tannenbaum, A. *Education of the Disadvantaged*. New York: Holt, Rinehart and Winston, pp. viii+503, £3.

Peters, M. L. *Spelling: Caught or Taught?* London: Routledge, pp. vii+88, 15s.

Price, K. *Education and Philosophical Thought*. Boston: Allyn and Bacon, pp. xiv+605, n.p.

Ramage, I. *Battle for the Free Mind*. London: Allen and Unwin, pp. 269, 40s.

Raven, J. C. *Psychological Principles Appropriate to Social and Clinical Problems*. London: Lewis (1966), pp. xx+187, 25s.

Richardson, E. *Group Study for Teachers*. London: Routledge, pp. xiii+130, 15s.

Rossel, G. *Manuel D'Éducation Psycho-Matrice*. Paris: Masson, pp. 156, 30F.

Ryle, A. *Neurosis in the Ordinary Family*. London: Tavistock, pp. 156, 32s. 6d.

Sand, E. A. *Contribution à l'Étude du Développement de l'Enfant.* Brussels : Université Libre, pp. 285, n.p.

Segal, S. S. *No Child is Ineducable.* Oxford : Pergamon Press, pp. xii + 332, n.p.

Smythies, J. R. (ed). *Science and E.S.P.* London : Routledge, pp. 306, 42s.

Spence, K. W. and J. T. (eds.). *The Psychology of Learning and Motivation*, Vol. 1. New York : Academic Press, pp. x + 381, 92s.

Slamecka, N. J. (ed.). *Human Learning and Memory—Selected Readings.* London : Oxford U.P., pp. x + 543, 46s.

Stephens, J. M. *The Process of Schooling.* New York : Holt, Rinehart, and Winston, pp. viii + 168, $3.95.

Tallent, N. *Psychological Perspectives on the Person.* London : Van Nostrand, pp. 292, 24s.

Tansley, A. E. *Reading and Remedial Reading.* London : Routledge, pp. xii + 176, 20s.

Trenaman, J. M. *Communication and Comprehension.* London : Longmans, pp. xx + 212, 35s.

Vail, E. *Tools of Teaching.* Illinois : Thomas, pp. ix + 163, n.p.

Warwick, D. W. *The Living Bible : Book* 4. London : Longmans, pp. 87, 7s. 6d.

Webb, L. *Children with Special Needs in the Infants' School.* London : Colin Smythe, pp. 216, 30s.

Weinberg, A. *Reflections on Big Science.* Cambridge, U.S.A. : M.I.T. Press, pp. ix + 181, 48s.

Weinberg, M. (ed). *School Integration—A Bibliography.* Chicago : Integrated Education Associates, pp. 137, $1.95.

West, D. J. *The Young Offender.* London : Duckworth, pp. 333, 35s. ; and Pelican Books, 6s.

Winters, G. *Musical Instruments in the Classroom.* London : Longmans, pp. 107, 10s. 6d.

Whalen, R. E. *Hormones and Behaviour.* London : Van Nostrand, pp. 266, 24s.

Wolfgang, M., and Ferracuti, F. *The Sub-Culture of Violence.* London : Tavistock, and New York : Barnes and Noble, pp. 387, 63s.

Zajonc, R. B. *Psychologie Sociale Expérimentale.* Paris : Dunod, pp. 139, n.p.

Enfance Aliénée. Numéro Spécial de Recherches. Paris : Bernard, pp. 318, NF20.

Mental Health Book Review Index, Vol. 12. New York : Council on Research in Bibliography, pp. xxv + 89, $8.00.

New Research in Education. London : Newnes, pp. 160, 15s.

Pedagogica Europaea (1965 and 1966). Edinburgh : Chambers, pp. 284 and 389, n.p.

Research Abstracts—Society for Research into Higher Education (July-Dec., 1966). Quarterly publication, pp. 50, annual subscription 40s.

The Scottish Standardisation of WISC. London : U.L.P., pp. 72, 11s. 6d.

Routledge & Kegan Paul

Experimental Psychology

Its Scope and Method

General Editors : PAUL FRAISSE and JEAN PIAGET

A nine-volume work which will be published in the course of three or four years. Each book covers the whole range of literature on the subject.

The first volumes to be published are :

Volume One

History and Method

PAUL FRAISSE, JEAN PIAGET and MAURICE REUCHLIN

Translated from the French by Judith Chambers

An account of the important stages in the growth of experimental psychology, from its origins in Aristotelian philosophy up to the most recent theories and experiments. 35s

Volume Five

Motivation, Emotion and Personality

JOSEPH NUTTIN, PAUL FRAISSE and RICHARD MEILI

Translated from the French by Betty Spillmann

Various kinds of motivation are considered, and their mechanism analysed. There is an examination of the nature of emotion, and of emotional situations and reactions, and the work ends with an account of the genesis and structure of personality, and the relation between its physical and psychological aspects. 35s.

Volume Seven

Intelligence

PIERRE OLERON, JEAN PIAGET, BARBEL INHELDER and PIERRE GRECO

Translated from the French by Thérèse Surridge

The first two chapters are on mental images and intellectual operations, followed by a chapter on intellectual acts and a description of relevant experiments. Learning and intellectual structures are considered and various interpretations are examined. About 35s.

AN INTRODUCTION TO EDUCATIONAL PSYCHOLOGY

E. Stones

"Here, at last, is the book we have been waiting for, an introduction to educational psychology which links theory to practice, and which draws on the most advanced thinking and recent scientific developments."
Forum

"This is a most welcome addition to the list of books on Psychology which can be recommended to students and teachers."
University of Newcastle Institute of Education Journal

"All branches of the subject are reviewed in sufficient detail to meet the needs of initial training courses . . . All in all the book fulfils its purpose admirably." *A.M.A.*

36s; University Paperback 18s

CHILD GUIDANCE AND THE SCHOOL

Isabella C. Maclean

"Dr. Maclean has produced a book that will meet a long-felt need among teachers coping with difficult children, and for many lecturers in education and psychology in departments and colleges of education."
The Teacher

Modern Teaching Series 18s; Education Paperback 9s 6d

INFANTILE AUTISM

B. Rimland

". . . it is essential reading for all who may be concerned in the care and study of abnormal children."
The Times Literary Supplement

"Undoubtedly the most important book ever written on the subject, this modestly priced text should be read by all paediatricians, psychiatrists, and psychologists."
Lector Medicus

35s

THE PSYCHOLOGY OF CHILDHOOD AND ADOLESCENCE

C. I. Sandström

"One can confidently recommend this book for reading, to teachers, educationalists and undergraduates taking their first steps in the study of developmental psychology, and also to those who are seeking to revise quickly and easily their knowledge of this subject."
Mental Welfare

Methuen's Manuals of Modern Psychology 36s

METHUEN

SOCIAL SCIENCE PAPERBACKS

Available in March

By the author of *The Child and the Family* and *The Child and the Outside World*

The Family and Individual Development

D. W. WINNICOTT

'These essays are packed with the author's very personal wisdom, humanity, and charm, and will illuminate the understanding of any adult who reads them.' *Archives of Disease in Childhood*

SSP 31 (Tavistock) **17s 6d net**
Hardback 30s net

Growth to Freedom

DEREK MILLER

'This book is warmly recommended to all those concerned with institutionalized children, particularly to those who are trying to find ways and means of helping such children to become more mature and self-reliant.' R. F. Barbour, *Journal of Child Psychology and Psychiatry*

SSP 30 (Tavistock) **15s net**
Hardback 30s net

Internal Factors in Evolution

LANCELOT LAW WHYTE

'The book is brief, trenchant and profound . . . All who are interested in the fundamental ideas of unity, ordering, complexity, adaptiveness and directiveness—and that is to say almost all biologists—should read this book.' W. H. Thorpe, *Nature*

SSP 32 (Tavistock) **12s 6d net**
Hardback 18s net

Estrangement and Relationship

FRANCIS A. MACNAB

'. . . this book is valuable for two different reasons. Firstly because it once again stresses the value of psychotherapy in the treatment of chronic schizophrenics. Secondly it is one of the few clear expositions of existential psychotherapy in the English literature.' F. Fish, *Medical News*

SSP 33 (Tavistock) **21s net**
Hardback 45s net

The Experimental Psychology of Beauty

C. W. VALENTINE

A comprehensive account and critique of experiments concerned with aesthetic appreciation and with the development of discrimination and 'taste' in painting, poetry, and music.

SSP 34 (Methuen) **25s net**
Hardback 50s net

Shakespeare: An Existential View

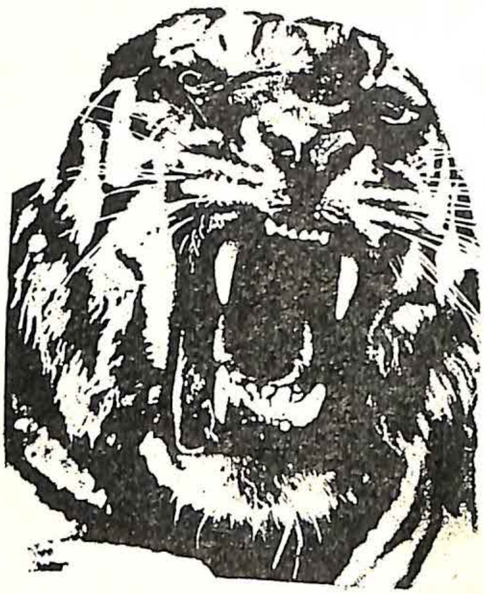
DAVID HOROWITZ

'. . . Mr. Horowitz has some real insight into the situations Shakespeare chooses to depict, and he relates the plays to our everyday concern with life with a refreshing directness.' Martin Dodsworth, *The Guardian*

SSP 28 (Tavistock) **15s net**
Hardback 25s net

Published by members of Associated Book Publishers Ltd.,
11, New Fetter Lane, London E.C.4.

Now in paperback



ON AGGRESSION

Konrad Lorenz

" A fascinating book by a master of his subject."

Sir Julian Huxley

" *On Aggression* is essential reading for all psychologists, psychoanalysts, sociologists and biologists—indeed for everyone who takes evolutionary theory seriously and who is concerned with the danger to civilisation implicit in our ignorance of the nature of aggression."

Charles Rycroft, *New Society*

" . . . the message is of profound significance for anyone concerned with the study of man."

J. M. Cullen, *Man*

University Paperback 10s 6d
Hard cover edition 30s

METHUEN

OCCUPATIONAL PSYCHOLOGY

Editor : ALEC RODGER

OCTOBER 1967

VOLUME 41, No. 4

CONTENTS

Communication Failure in Questioning: Errors, Misinterpretations and Personal Frames of Reference. By MARY SPEAK
Individualism, Home Life and Work Efficiency Among a Group of Nigerian Workers. By MALLORY WOBER
A Study of and Some Concepts Concerning Labor Turnover. By W. W. RONAN
Describing Non-Repetitive Tasks for Training Purposes. By D. J. CUNNINGHAM and K. D. DUNCAN
Task Analysis and Training Design. By J. ANNETT and K. D. DUNCAN
The Vocational Aspirations of Educationally Subnormal Children. By K. LAWSON and C. J. HARTLEY
The Managerial Grid. By MAX BROLLY
Age and Re-Employment: A Regional Study of External Mobility. By J. M. SMITH
The Partial Validation of a Leaderless Group Discussion for the Selection of Supervisory Personnel. By C. L. JAFFEE.
Personality Patterns in Various Groups of Businessmen. By H. J. EYSENCK
Book Reviews

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : R. TAFT, University of Melbourne
Associate Editor : R. H. DAY, Monash University.

August, 1967

Vol. 19 *No.* 2

DAVENPORT, W. G. and HALFORD, G. S. Incremental and All-or-None Learning with a Stimulus Complex.

COLLINS, J. K. A Kinesthetic Spatial Aftereffect with Pronation of the Forearm.

TAUSS, W. A note on the Prevalence of Mental Disturbance.

FRANKS, CYRIL M. A Longitudinal Conditional Eyeblink Reflex Study of Four Adult Men.

KRAUS, J., CHALKER, SUZANNE, and MACINDOE, I. Vocabulary and Chronological Age as Predictors of " Abstraction " on the Shipley-Hartford Retreat Scale.

PROVINS, K. A. Motor Skills, Handedness and Behaviour.

SARBIN, THEODORE R., and KI-TAEK CHUN. A Logical Flaw in an Index to Suppress Response-Bias in Perceptual Defence Measures and the Application of a Proposed Improvement.

FUREDY, J. J. Aspects of Reinforcement Through UCS Offset in Classical Aversive Conditioning.

The *Journal* is published three times a year. Subscription rate $A5.50. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to The Executive Officer, Australian Psychological Society, Department of Psychology, University of Sydney, Sydney, 2006, Australia.

CAMBRIDGE

The British Journal of Psychology

Edited by **W. SLUCKIN**
Professor of Psychology, University of Leicester

The British Journal of Psychology is a journal of general and experimental psychology which was founded in 1904. It is particularly concerned with human psychology and covers methods of testing for and measuring psychological variants, personality assessment and aptitude measurement.

A series of monograph supplements was issued in connection with the journal. Details of these will be sent on request to the publishers.

Annual subscription 90s. net ($16.50)
Single parts 30s. net ($6.50)

CAMBRIDGE UNIVERSITY PRESS

SOCIOLOGICAL REVIEW MONOGRAPH No. 12

Monograph Editor : Paul Halmos

"The Foundation Year in the University of Keele"

a report by

A. H. ILIFFE

Nuffield Research Fellow 1962-64

December, 1967

Price : 30s. +11d. postage
U.S. : $4.60 +20c postage

Copies are obtainable from : The Secretary to the Monograph Editor,
Sociological Review Monographs, University of Keele, KEELE, Staffordshire,
ENGLAND.

The British Journal of Medical Psychology

Volume 41, Part I. March 1968

Proceedings of Symposium on Research in Psychotherapy.

I. R. C. Batchelor Introduction

Jerome D. Frank Recent American Research in Psychotherapy

J. M. M. Mair and A. H. Crisp Estimating Psychological Organisation, Meaning and Change in Relation to Clinical Practice

Sidney Crown Criteria for the Measurement of Outcome in Psychotherapy

M. G. Gelder Desensitisation and Psychotherapy Research.

I. M. Marks Aversion Therapy

Thomas Freeman A Psycho-analytic Critique of Behaviour Therapy

H. J. Walton and F. M. McPherson Phenomena in a Closed Psychotherapeutic Group

D. Russell Davis Interventions into Family Affairs

30s. net ($6.50). Annual subscription 90s. net ($16.50).

Orders should be sent to a bookseller or to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1, England, or Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y.10022, U.S.A.

The British Journal of Social and Clinical Psychology

Volume 6, Part 4. December, 1967

E. Clifford & M. Clifford Self-concepts before and after survival training

N. P. Pollis Relative stability of scales formed in individual, togetherness and group situations

R. V. Exline & D. Messick The effects of dependency and social reinforcement upon visual behaviour during an interview.

P. B. Warr, J. Faust & G. J. Harrison A British ethnocentrism scale

M. Hamilton Development of a rating scale for primary depressive illness

E. Blundell A psychological study of the effects of surgery on eighty-six elderly patients

A. W. Heim, K. P. Watts & V. Simmonds The Brook reaction as a test of temperament

30s. net ($5.50).

Annual Subscription 90s. net ($16.50) for four issues.

Orders should be sent to bookseller or to Cambridge University Press, Bentley House, P.O. Box 92, 200 Euston Road, London, N.W.1, England, or Cambridge University Press, American Branch, 32, East 57th Street, New York, N.Y.10022, U.S.A.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (*Monograph Supplements*).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor W. SLUCKIN, Department of Psychology, The University, Leicester.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Professor J. D. NISBET, King's College, Aberdeen.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. RALPH HETHERINGTON (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor IAN M. L. HUNTER, Department of Psychology, University of Keele, Keele, Staffs., whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

25 MAR 1969

CONTENTS

H. B. GIBSON: Parental Attitudes and their Relation to Boys' Behaviour . 233
K. LOVELL and A. GORTON: Some Differences between Backward and Normal Readers of Average Intelligence 240
D. JOYNT and B. CAMBOURNE: Psycholinguistic Development and the Control of Behaviour 249
E. STONES and J. R. HESLOP: The Formation and Extension of Class Concepts in Primary School Children 261
H. J. BUTCHER and H. B. PONT: Opinions about Careers among Scottish Secondary School Children of High Ability 272
J. H. CLARKE: The Image of the Teacher 280
P. K. POPPLETON: Puberty, Family Size and the Educational Progress of Girls 286
B. P. FROST: Anxiety and Educational Achievement . . . 293
A. J. PICKUP and W. S. ANTHONY: Teachers' Marks and Pupils' Expectations . . . 302

RESEARCH NOTES:

S. LEMESHOW: A Problem Check List for Teachers 310
C. K. MACKAY and T. M. KILKENNY: The Stability of Induced Conservation 313
C. K. MACKAY and M. B. CAMERON: Cognitive Bias in Scottish First-year Science and Arts Undergraduates 315
F. FEE: An Alternative to Ward's Factor Analysis of Wallach and Kogan's 'Creativity' Correlations (Comment by J. Ward) 319

THESES ABSTRACTS:

R. HUME: The Effect of School and Curriculum on Boys' Attitudes towards Careers in Engineering 322
R. AMOS: Teachers' Opinions on the Place of Scientific Method in Biology Courses 323
C. L. BAIRD: The Role of the Teacher of Six- and Seven-year-old Children 323
B. STAVELEY: The Abilities and Interests of Craft and Technician Students of Mechanical Engineering 324
BOOK REVIEWS 325
PUBLICATIONS RECEIVED 336
INDEX TO VOLUME 38 338
NOTES TO CONTRIBUTORS 341

VOLUME 38 JUNE 1968 PART 2

PREDICTING STUDENT PERFORMANCE IN COLLEGES OF EDUCATION

BY G. A. CORTIS*

(*Furzedown College, London*)

SUMMARY. Two-hundred-and-fifty-nine students in three constituent colleges of education of the University of Manchester School of Education were tested with cognitive, creativity and personality measures. They also supplied some biographical details. The predictive power of these measures was analysed in terms of final examination grades in educational theory, practical teaching, academic subjects and basic English. The analysis revealed an apparent dichotomy between the professional and academic criteria of success. Correlates of classroom success included previous teaching experience, lower verbal fluency and naturalistic and tender-minded attitudes to education. Correlates of academic success included high verbal ability, sound entrance qualifications and idealistic and tough-minded attitudes to education. The results obtained are compared with a similar investigation of students from the University's Department of Education.

I.—OUTLINE OF PREVIOUS RESEARCH.

As Warburton, Butcher and Forrest (1963, p. 68) note, the "literature dealing with the prediction of success in teaching and success in teachers' training courses is very large and very inconclusive." Many variables have been studied in relationship to teaching and academic success: intelligence and verbal abilities (Vernon, 1939; Shea, 1955; Warburton, 1955; Tarpey, 1965), creativity (Yamamoto, 1963), educational attitudes (Evans, 1953; Oliver and Butcher, 1962), personality factors from both questionnaire and objective-type tests (Warburton and Hadley, 1960; Soloman, 1965) and biographical data such as extra-curricular activities and sex differences (Lawton, 1939; Evans, 1957). The validity of interview-type selection methods has also been extensively examined (Burroughs, 1958; Drever, 1963). None of these studies has given a firm answer to prediction problems in this sphere and this uncertainty arises not only from the complex nature of the criterion, but equally from the different nature of the samples and the various types of measure used.

This present investigation, by using a number of the same measures as the university sample of Warburton, Butcher and Forrest (1963), provides useful comparisons of the 'best' predictors in each group. These comparisons, however, are inevitably limited both by the different academic background of the two samples and by the different criteria employed.

II.—PLAN OF THE INVESTIGATION.

(1) *The Sample.* Three colleges of education in the University of Manchester School of Education were used for the experiment. Two were LEA colleges, one a voluntary college. The subjects were students completing their three-year training courses for the Teacher's Certificate in June, 1965. Over 300 students began the five-hour test battery but inevitably, at the completion of the battery, the numbers were reduced. Finally, effective information was obtained on 259 students (152 women, 107 men). Of this number 132 were 'volunteers' and 127 were 'non-volunteers' who took the test battery as part of an 'end-of-course' exercise. There were considerable differences in age, social background

---

* Formerly at Mather College, Manchester.

and attainment amongst the sample although its proportions (women to men, local authority to voluntary colleges) bore no statistical resemblance to any national pattern of college populations.

(2) *The Test Battery*. Three of the tests used are well-known, well-standardised and need no comment (Heim's AH5, Cattell's 16 PF Questionnaire, Oliver's 'A Survey of Opinions about Education'). Brief notes are appended on the others as follows:

(*a*) *Vocabulary*. This was an adaptation of Thurstone's Primary Mental Abilities test used by Warburton, Butcher and Forrest (1963). No reliability or validity figures are available for this adaptation.

(*b*) *Culture*. Richmond's (1963) Test 2 was used. This contains forty multiple-choice items and gives scores for Arts and Science Culture categories. The author indicates it is highly reliable but its validity is more difficult to assess.

(*c*) *Creativity*. Three tests of Getzels and Jackson (1962) were used: Word Association, Uses for Objects, Fables. Slight variations were made in the scoring methods in this present investigation. Reliability and validity coefficients in both cases are uncertain, for these tests are still at an experimental stage.

(*d*) *Objective Personality Tests*. These were adaptations of Cattell's Objective-Analytic battery. Some are questionnaire-type items, such as estimating level of performance in untried tasks. Others are performance-type items, such as copying a series of numbers in normal sequence against a time limit and then writing them backwards against an even more severe time limit. The reliability and validity of the adapted models are uncertain but Cattell (1965, p. 127) makes strong claims for the originals.

(3) *The Criteria*. These were the Teacher's Certificate Examination final results which had been set centrally by the university, examined internally within colleges and moderated by external examiners nominated by the university.

(4) *The Variables*. These comprised forty-seven test scores, twelve biographical details and six criteria of course-success as in the Appendix.

(5) *Analysis of Data*. Means, standard deviations and product moment correlations between the sixty-five variables were calculated on the Manchester University electronic computer. The resultant 2,080 correlations were subjected to two factor analyses—a Principal Components solution and one by the Varimax method. The latter method is discussed here since this tends to spread loadings more evenly between a number of group factors than does the former (Harman, 1965).

III.—RESULTS.

(*a*) *Intercorrelations*. (Correlations in Table 1 are significant at the 5 per cent level unless asterisked when they are significant at the 1 per cent level. Non-significant correlations are omitted. Decimal points are omitted.)

In the cognitive area a similar percentage (19) of significant correlations to that of the university sample was obtained. The pattern is similar too, except that verbal intelligence shows a positive and significant relationship to the criteria. This might be due to the college sample having a wider intellectual range than the university sample or to random variation arising from the different intelligence tests used. Grades on practical teaching did not correlate

TABLE 1

SIGNIFICANT CORRELATIONS BETWEEN PREDICTORS AND CRITERIA OF SUCCESS.

Predictors	Criteria					
	Educational Theory	Practical Teaching	Main Subject	Subsidiary Subject 1	Subsidiary Subject 2	Basic English
COGNITIVE:						
1—Vocabulary	191*		219*		126	255*
2—Verbal Intelligence			239*	144		190*
4—Culture: Arts						258*
6—Creativity: Word Association	148		220*			
EDUCATIONAL ATITUDES:						
10—Naturalism in Education		182*				
12—Tender-mindedness in Education	137				149	152
16 PERSONALITY FACTOR QUESTIONNAIRE:						
13—Friendliness (Factor A)						143
14—Intelligence (Factor B)				144	204*	140
20—Sensitivity (Factor I)	240*					173*
21—Suspiciousness (Factor L)					−194*	
22—Unconventionality (Factor M)	125					
25—Radicalism (Factor Q1)						123
OBJECTIVE PERSONALITY TESTS:						
33—Essay: more adjectives			−140			169*
37—More friends named						132
38—Limited experience paired with more skill				124		
39—Backward Writing: superior performance	−138			−216*		−211*
40—Addition: higher score	−161			−255*	−131	−184*
41—Higher annoyability						−148
42—More 'unconventional' names chosen				137		
44—Higher self-rated anxiety	124					126
45—Higher self-rated stability	−164*					−212*
47—Volunteering: more team tests			−170*		−195*	−142
BIOGRAPHICAL DATA:						
51—Number of G.C.E. 'O'	187*			242*	183*	261*
52—Number of G.C.E. 'A'	246*		416*	222*	189*	
55—Number of interests						140
56—Number of games						−151

significantly with any of the cognitive variables, which may indicate that teaching skill bears no direct relationship to cognitive ability.

Among the correlations between educational attitudes and criteria appears the only significant correlation of any of the predictor variables with practical teaching, namely naturalism in education, $+0{\cdot}182$ ($p < {\cdot}01$). In the university sample, the correlation was in the opposite direction (towards idealism in education) and away from the progressive, child-centred grouping of attitudes which naturalism in education represents (Butcher 1959; Oliver and Butcher 1962).

The percentage of significant correlations in the educational attitudes area was the same in both samples (22 per cent).

The significant correlations of the 16 PF personality questionnaire scores with criteria were relatively few, totalling only 8 per cent of significant correlations compared with the university sample's 18 per cent. Two of the correlations, sensitivity and unconventionality, bear the same positive relationship to the educational theory examination as in the university sample. Unlike that sample, however, the 16 PF has failed to reveal any significant relationships to practical teaching such as Warburton, Butcher and Forrest (1963) found in the case of conscientiousness, sensitivity and self-discipline.

The correlations of objective personality tests and criteria embrace a new area in terms of test material with student teachers. Solomon (1965) has recently reported some work on these measures confined to mature students exclusively. Twenty of the 102 correlations in this area are significant. This is almost 20 per cent and compares favourably with more well-established tests. In general, the correlations associate success in written examinations with the following variables of the test battery: 37 friendship, 38 higher confidence in attempted skill, 39 and 40 poorer performance in stress conditions (negative correlations with better performance), 41 less annoyability (negative correlation), 42 unconventionality, 44 and 45 higher self-rated anxiety (one positive with anxiety, one negative with stability), 47 individuality/self-sufficiency (negative correlation with team volunteering).

Finally, of the correlations between biographical data and criteria, ten of the seventy-two were significant, 14 per cent compared to the 25 per cent of the university sample. They were largely in the expected direction, the coefficient between number of 'A' levels and main subject (·416) being the highest correlation of any of the predictor variables used in the experiment.

The total number of correlations between all predictor variables and criteria was 354 of which fifty-three were significant (15 per cent). Of the 282 correlations between test predictors alone and criteria, thirty-three were significant. This is 11·7 per cent compared with the university sample's 13 per cent in the same areas.

(*b*) *Factor analysis.* The factor analysis revealed nineteen major factors when rotated to the Varimax Criterion and these were interpreted as in Table 2. The twelve factors with criterion loadings are further analysed in Table 3.

TABLE 2

FACTORS EXTRACTED BY VARIMAX ROTATION.*

Factors without criterion loadings	Factors with criterion loadings
1—Stability.	4—College Success.
2—Extraversion.	5—Intellectual-Science Bias.
3—Mature Student Entry.	6—Verbal Fluency.
7—Female Aesthetic Interests.	10—Family Size.
8—Sociability.	11—Educational Attitudes.
9—Verbal Creativity.	13—Social Confidence.
12—Self-Sufficiency.	14—Criticalness.
	15—Venturesomeness.
	16—Teaching Experience.
	17—Radicalism.
	18—Verbal Extraversion.
	19—Unconventionality.

* In descending order of peercentage variance.

TABLE 3

VARIMAX FACTORS WITH LOADINGS ON CRITERION VARIABLES.

Examination	Number and Name of factor and relevant loadings on examination success	
Educational Theory	18—Radicalism	.211
Practical Teaching	17—Previous Teaching Experience	·294
	6—Lower Verbal Fluency	·219
	10—Large Family Size	·187
	16—Venturesomeness	·186
	11—Naturalist and Tender-minded Educational Attitudes	·163
Main Subject	5—Intellectual-Science Bias	·222
Subsidiary Subjects and Basic English	6—Verbal Fluency	·319
	15—Less Criticalness	·245
	17—No Previous Teaching Experience	·190
	19—Verbal Introversion	·183
	11—Idealistic and Tough-minded Educational Attitudes	·181
	20—Conventionality	·171
	13—Social Confidence	·166

IV.—DISCUSSION.

Warburton, Butcher and Forrest (1963, p. 74), in discussing their intercorrelations, indicate the familiar problem of correction for range. They state " There is little doubt that if the tests and predictors had been administered to all candidates, the majority of correlation coefficients would have been markedly higher." Some recent comments about validity have led to uncertainty in the interpretation of predictive studies. Drever (1963, p. 410) notes " What a particular coefficient means for practical purposes will depend on the selection ratio. As a rough guide, a correlation of +0·2 might be significant but it is too low to be useful." Oliver (1962, p. 265), discussing the relationship of GCE 'A' level results and degree class, notes that they correlate " about 0·4. This means that selecting university students on the basis of GCE results is better than selecting them from a list of names with a pin, but not vastly better." Since many of the correlation coefficients (and factor loadings) are in this relatively low range, both here and in the university study, these seem damning comments. Both Oliver and Drever, however, are referring to the criterion of performance in degree examinations. Where selection is being evaluated in terms of occupational success, the degree of validity " depends on the type of decision-situation in which it (the test) is applied. Thus, in some circumstances interview judgments with a validity of 0·2 may be preferable to intelligence tests with a validity of 0·7 " (Vernon, 1958, p. 81).

Bearing in mind then that the correlation coefficients and factor loadings amongst the college sample range from ·123 to ·416 and are, as in many predictive studies, on the low side the following variables appear related to success in the training course :

(*a*) Amongst the more successful students in practical teaching, previous teaching experience, lower verbal fluency, membership of a larger family and naturalistic and tender-minded attitudes to education are correlates of success.

(*b*) Correlates of academic success include high verbal ability, sound entrance qualifications ('A' levels in particular), no previous teaching experience and idealistic and tough-minded attitudes to education. In the case of educational theory radical attitudes to life and education, together with sensitivity are related to success.

Considering the test areas common to both samples, the university student appears to be less ' child-centred ' (attitudes to education test) and to show better mental health in the classroom (16 PF factors of conscientiousness sensitivity and self-discipline). Whether that difference arises from the different nature of the schools in which the teaching was done is a matter for conjecture. A recent investigation (Evans, 1964) of two samples of university and college students concluded that the two groups were equally ' stable ' on personality tests. So it is certainly possible that had the university sample here been removed from their more favourable teaching milieu to that of some of the colleges' urban schools (in areas of low socio-economic status), the present positive correlations might well have been lower, if not negative. Both the Newsom and Plowden reports have noted the difficulties of teaching in schools of this type.

In the college sample, the absence of criterion loadings on factors such as stability and self-sufficiency, and the association, for example, of lower verbal fluency with teaching success raise interesting questions as to the nature of the sample and the criteria used. If two distinct areas exist within colleges of education, the professional and the academic, each having different criteria of success (and only further extensive research will demonstrate whether this is so nationally), then the design of any equable selection procedure or its evaluation will be difficult. Particularly in view of the impending award of B.Ed. degrees this apparent dichotomy would seem to warrant further investigation.

V.—References.

Burroughs, G. E. R. (1958). A study of the interview in the selection of students for teacher training. *Brit. J. Educ. Psychol.*, 28, 37-46.

Butcher, H. J. (1959). *The Attitudes of Teachers and Student Teachers to Education.* Unpublished Ph.D. thesis, University of Manchester.

Cattell, R. A. (1965). *The Scientific Analysis of Personality.* London : Penguin Books.

Cortis, G. A. (1966). *The Prediction of Student Performance in Colleges of Education.* Unpublished M.Ed. thesis, University of Manchester.

Drever, J. (1963). Prediction, placement and choice in university selection. *Higher Education*, Appendix 2 (B), London : H.M.S.O.

Evans, E. G. S. (1964). Reasoning ability and personality differences among student teachers. *Brit. J. Educ. Psychol.*, 34, 305-314.

Evans, K. M. (1953). A further study of attitude towards teaching as a career. *Brit. J. Educ. Psychol.*, 23, 58-63.

Evans, K. M. (1957). Is the concept of ' Interest ' of significance to success in a teacher training course ? *Educ. Rev.*, 9, 205-211.

Getzels, J. W., and Jackson, P. W. (1962). *Creativity and Intelligence.* London and New York : John Wiley.

Harman, H. H. (1965). *Modern Factor Analysis.* Chicago : Univ. Chicago Press.

Lawton, J. A. (1939). A study of factors useful in choosing candidates for the teaching profession. *Brit. J. Educ. Psychol.*, 9, 131-144.

Oliver, R. A. C. (1962). The selection of university students : a ' scholastic aptitude test ' ? *Univ. Quart.*, 16, 264-273.

Oliver, R. A. C., and Butcher, H. J. (1962). Teachers' attitudes to education. *Brit. J. Soc. Clin., Psychol.*, 1, 56-59.

Richmond, W. Kenneth (1963). *Culture and General Education.* London : Methuen.

Shea, J. A. (1955). The predictive value of various combinations of standardised tests and sub-tests for prognosis of teaching efficiency. *Cath. Univ. of Amer. Res. Monogr.*, 19, 6.

Solomon, E. (1965). *Personality Factors and Attitudes of Mature Training College Students.* Unpublished M.Ed. thesis, University of Manchester.

TARPEY, M. S. (1965). Personality factors in teacher trainee selection. *Brit. J. Educ. Psychol.*, 35, 140-149.

VERNON, P. E. (1939). Educational abilities of training college students. *Brit. J. Educ. Psychol.*, 9, 233-250.

VERNON, P. E. (1958). Critical Notice. "Cronbach, L. J., and Gleser, G. C.—Psychological Tests and Personnel Decisions." *Brit. J. Educ. Psychol.*, 28, 80-83.

WARBURTON, F. W. (1955). *The Selection of Students in a University Department of Education.* London: Organisation mondiale pour l'education pre-scolaire.

WARBURTON, F. W., and HADLEY, S. T. (1960). Predicting the achievement in teacher training courses of college students studying to be teachers. *I.P.A.T. Information Bulletin* 4. Champaign, Illinois; Institute for Personality and Ability Testing.

WARBURTON, F. W., BUTCHER, H. J., and FORREST, G. M. (1963). Predicting student performance in a university department of education. *Brit. J. Educ. Psychol.*, 33, 68-79.

YAMAMOTO, K. (1963). Relationships between creative thinking abilities of teachers and adjustment of pupils. *J. Exp. Educ.*, 32, 3-25.

VI.—Appendix.

LIST OF EXPERIMENTAL VARIABLES

A.—Cognitive Tests.

1—Vocabulary: Adaptation of Primary Mental Abilities (Thurstone).
2—AH5 (Heim): Verbal intelligence.
3—AH5 (Heim): Non-verbal intelligence.
4—Culture (Richmond): Arts.
5—Culture (Richmond): Science
Creativity (Getzels and Jackson); 6—Word Association; 7—Uses for Objects; 8—Fables (Appropriateness): 9—Fables (Relatedness).

B.—Educational Attitudes Test.

Survey of Opinions about Education (Oliver), Department of Education, Manchester University.

10—Naturalist attitudes to education.
11—Radical attitudes to education.
12—Tender-minded attitudes to education.

C.—Personality Questionnaire.

16 Personality Factor Scale (Cattell)—(Name of positive pole given).

13—Friendliness (A).
14—Intelligence (B).
15—Stability (C).
16—Dominance (E).
17—Surgency (F).
18—Conscientiousness (G).
19—Unreservedness (H).
20—Sensitivity (I).
21—Suspiciousness (L).
22—Unconventionality (M).
23—Shrewdness (N).
24—Insecurity (O).
25—Radicalism (Q^1).
26—Self-Sufficiency (Q^2).
27—Self-Discipline (Q^3).
28—Tenseness (Q^4).
29—Second Order Factor: Extraversion.
30—Second Order Factor: Anxiety.

D.—Objective Personality Tests.

31—Goodness of work: subject rates various work performances.
32—Estimation of Time: subject rates his own and other people's performances at various tasks.
33—Essay: number of adjectives used as proportion of total words.
34—Essay: number of verbs used as proportion of total words.
35—Crime and Punishment: matching of offences with punishments.
36—Friends and Acquaintances: number of friends plus acquaintances named.

37—Friends and Acquaintances: numerical difference between friends and acquaintances named.
38—Survey of Experience and Skills: pairing personal experience with various skills.
39—Backward Writing: writing a word forwards then backwards against time limits.
40—Addition: adding '1' to a series of figures against time limits.
41—Common Annoyances: rating annoying situations.
42—Alternative Names: picking pairs of 'unconventional' names from sets of five.
43—Innovations: rating innovations in modern living.
44—Test Comments: self-rating for anxiety.
45—Test Comments: self-rating for stability.
46—Volunteering: number of tests volunteered for as individual.
47—Volunteering: number of tests volunteered for as team member.

E.—BIOGRAPHICAL DETAILS.

48—Age.
49—6th Form Experience: duration.
50—School-leaving Age.
51—Number of G.C.E. 'O' level subjects passed.
52—Number of G.C.E. 'A' level subjects passed.
53—Pre-College Employment: duration.
54—Previous Teaching Experience: duration of paid service.
55—Number of interests stated.
56—Number of games stated.
57—Social class of father's occupation: Grades I–V (Classification of Occupations, H.M.S.O. 1960).
58—Family size: number of children in student's family.
59—Student's position in family: first-born, second-born, etc.

F.—CRITERIA OF SUCCESS.

Final Grades in University School of Education Examinations A–E Scale.
60—Grade in educational theory.
61—Grade in practical teaching.
62—Grade in academic main subject.
63—Grade in first subsidiary academic subject.
64—Grade in second subsidiary academic subject.
65—Grade in basic English.

(*Manuscript received* 3*rd May*, 1967)

NEUROTICISM AND SCHOOL ATTAINMENT—A LINEAR RELATIONSHIP?

By N. J. ENTWISTLE* and SHIRLEY CUNNINGHAM
(*Department of Education, Aberdeen University*)

Summary. In a follow-up study 2,995 Aberdeen children aged about 13 years were given the Junior Eysenck Personality Inventory. School attainment was measured by teachers' average rank order in class, scaled on a verbal reasoning test. The relationships between attainment and the personality dimensions of neuroticism and extraversion were examined. Evidence is presented which suggests that the relationship between neuroticism and attainment is linear; high attainment in this age-group was associated with stability. The results for extraversion showed a sex difference, which may explain earlier conflicting findings. Girls who were stable extraverts and boys who were stable introverts formed superior groups. These relationships were also present among the brightest children.

I.—Introduction.

The relationships between academic attainment and scores on the personality inventories devised by Eysenck and Cattell have been investigated for both students and school children. The findings for these two populations show distinct differences.

At university level, Furneaux (1962) reported that a group of 'neurotic introverts' had the lowest failure rate in examinations. 'Stable extraverts' had the highest failure rate. Kelvin and his co-workers (1965) confirmed the superiority of the 'neurotic introverts,' but found that a group of student failures tended to be 'neurotic extraverts.' Also at university level, Lynn and Gordon (1961) suggested that the Yerkes-Dodson 'law' might indicate a non-linear relationship between anxiety or neuroticism and intelligence test scores, both high and low anxiety being associated with below average scores. They showed a regression curve of matrices score on neuroticism, which was of the expected inverted-U shape. But this curve was derived from a total group of only sixty students.

Among school children different results have been obtained. Rushton (1966), using Cattell's Children's Personality Questionnaire, found that the second order factors of both anxiety and neuroticism correlated negatively with verbal reasoning score and school record, whereas extraversion showed a positive correlation with school attainment. Again, Hallworth (1964) showed that teachers apparently base their personality ratings on the two main dimensions of social extraversion and emotional stability. Both these dimensions showed a positive correlation with school attainment. A conflicting finding was reported by Child (1964), using the Junior Maudsley scale. He found that stable introversion was related to high achievement in a sample which was not strictly representative.

Correlational studies assume linear relationships between the variables being investigated. Eysenck and White (1964) suggested that the low values of the correlation coefficients between personality and intellectual measures in many studies might be caused by a non-linearity of relationship, similar to that found by Lynn and Gordon (1961). In the same article, Eysenck and White presented evidence which led them to suggest that children of different personality types may also have different intellectual structures. It might thus be possible to find different relationships between intelligence and personality, when different types of test were used. There might also be different relationships between these variables at different intellectual levels, which, in turn, might explain the differences between results obtained with university students and with school children.

---

* Now at Department of Educational Research, University of Lancaster.

The present study is based on data collected as part of an investigation into the age of transfer to secondary education in Scotland. The initial stages of this longitudinal study have already been reported (Nisbet and Entwistle, 1966). The availability of JEPI scores and intellectual test scores on a nearly complete age-group of 13-year-olds suggested this digression from the main purpose of the follow-up study.

The previous results from studies of school children raised the following questions. These also formed the hypotheses to be tested. (1) Is school attainment negatively related to neuroticism and positively related to extraversion, as measured by the JEPI? But (2) are the true relationships both non-linear? (3) Is there any sex difference in these relationships? (4) At the higher levels of attainment, do the 'neurotic introverts' form a superior group compared with other personality types?

II.—METHOD.

The original age-group tested in 1964 consisted of 3,286 children, who were attending schools in the city of Aberdeen. Of these children 2,995 took the personality test in 1966; there are complete sets of data for 2,707 children between the ages of 7 and 13 years. The correlational analysis, reported later, is based on the depleted group, whereas the regression analysis is based on the larger group. The following measures have been included in these analyses.

At age 11 years: NFER Non-verbal test 1.

At age 12 years: Moray House Verbal Reasoning test 72.

At age 12 years: Teachers' estimates of English and arithmetic attainment in the primary school, scaled against Moray House attainment tests. This method of scaling imposed on the teachers' average rank-order for each class the mean score and standard deviation obtained by that class in the appropriate standardised test. The final measure averaged attainment in English and arithmetic (see Yates and Pidgeon, 1957, page 88).

At age 13 years: Junior Eysenck Personality Inventory.

At age 13 years: Secondary school teachers' average rank-orders in all academic subjects, scaled on Moray House Advanced Verbal Reasoning test 12.

The intellectual tests were restandardised to a mean of 100 and standard deviation of 15 on the Aberdeen age-group. The internal reliability coefficients of the intellectual tests are all higher than ·90 and their predictive validities have been repeatedly demonstrated by their use in selection batteries. The JEPI scales have internal reliability coefficients of about ·80 and test-retest reliability coefficients of about ·70 at age 13 (Sybil Eysenck, 1965b). The validity of this instrument has not been clearly demonstrated, except on clinical subjects with extreme symptoms of neuroticism and extraversion.

III.—RESULTS.

The analysis in section (A) opposite is based on the depleted group of 2,707 children who had complete sets of test scores. In the following analyses 'significance' implies a probability level of less than ·05.

(A) *Means and Correlation Coefficients.*

TABLE 1

MEANS AND STANDARD DEVIATIONS OF N, E AND L SCORES BY SEX.

	Boys		Girls		Total	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Neuroticism	11·832	(4·735)	12·910	(4·432)	12·358	(4·763)
Extraversion	17·858	(3·748)	17·224	(3·620	17·548	(3·699)
Lie Scale ..	4·154	(2·217)	4·531	(2·359)	4·338	(2·295)
(N)	(1385)		(1322)		(2707)	

Table 1 shows values similar to those reported by Sybil Eysenck (1965b) in the JEPI manual. At age 13 years the boys tend to be more extraverted than the girls and also more stable.

TABLE 2

INTER-CORRELATION COEFFICIENTS BETWEEN THE PERSONALITY SCALES.

Scale	Boys		Girls		Total	
	N	E	N	E	N	E
Extraversion	—·221	—	—·173	—	—·206	—
Lie Scale ..	—·261	—·068	—·254	—·103	—·246	—·091

The inter-relationships between the scores derived from the JEPI (shown in Table 2) are also similar to those shown in the manual. The relationship between neuroticism and extraversion is not strictly orthogonal; the values of the correlation coefficients for both boys and girls are significantly above zero. Although these values are not high, there is still a tendency for the neurotic children to be introverted and for the stable children to be extraverted.

TABLE 3

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN NEUROTICISM AND INTELLECTUAL SCORES AT VARIOUS AGES BY SEX.

Test	Age	Boys	Girls	Total
Teachers' Estimates	13	—·137	—·185	—·157
Verbal Reasoning	13	—·129	—·158	—·141
Teachers' Estimates	12	—·120	—·163	—·135
Verbal Reasoning	12	—·127	—·143	—·134
Non-Verbal	11	—·119	—·128	—·123
(N)		(1385)	(1322)	(2707)

TABLE 4

PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN EXTRAVERSION AND INTELLECTUAL SCORES AT VARIOUS AGES BY SEX.

Test	Age	Boys	Girls	Total
Teachers' Estimates	13	—·021	·064	·018
Verbal Reasoning	13	·018	·084	·048
Teachers' Estimates	12	—·001	·083	·035
Verbal Reasoning	12	·022	·080	·050
Non-Verbal	11	·035	·083	·058
(N)		(1385)	(1322)	(2707)

The starting point of this investigation was the question " Is school attainment negatively related to neuroticism and positively related to extraversion ? " Tables 3 and 4 show the values of the correlation coefficients obtained. As the results from the various intellectual tests at different ages show no significant differences, the statistical analysis was limited to the measure of attainment at age 13 years.

The results shown in Table 3 support the previous findings. School attainment and neuroticism are negatively related ; children with low neuroticism scores tend to occupy higher positions in class than those with high neuroticism scores. The relationship is by no means close, but in this large representative sample it is statistically significant. The relationship is slightly closer for girls than for boys, but this sex difference could have arisen by chance.

Table 4 presents a different pattern of relationships. The correlations between extraversion and school attainment are small and could have occurred by chance. There is, however, an interesting sign difference between the correlation coefficients obtained for boys and for girls. This sex difference is statistically significant and will be shown by the following regression analysis to be distinct.

(B) *Regression Analysis.*

The next stage was to test the second hypothesis. Are the relationships between the personality dimensions and school attainment non-linear? Previous studies suggested that they might well be non-linear. A regression analysis was applied to data from all 2,995 children who completed the personality inventory, but it was limited to a single attainment measure, namely the teachers' estimates at age 13 years. Mean attainment quotients at five different levels of neuroticism and four different levels of extraversion were calculated, keeping the sexes separate. This allowed the variation of attainment with neuroticism or extraversion to be plotted graphically. Linearity or non-linearity shows up clearly using this approach. An analysis of variance method (McNemar, 1962, pages 275-281) was used to test whether the relationships were significantly non-linear.

Figure 1 suggests that the relationship between neuroticism and attainment is approximately linear. The analysis of variance showed no significant divergence from linearity (F= ·99 with d.f. of 2990 and 3). Again the relationship is similar for boys and girls. The hypothesis of non-linearity and in particular of an inverted-U relationship must be emphatically rejected on this evidence.

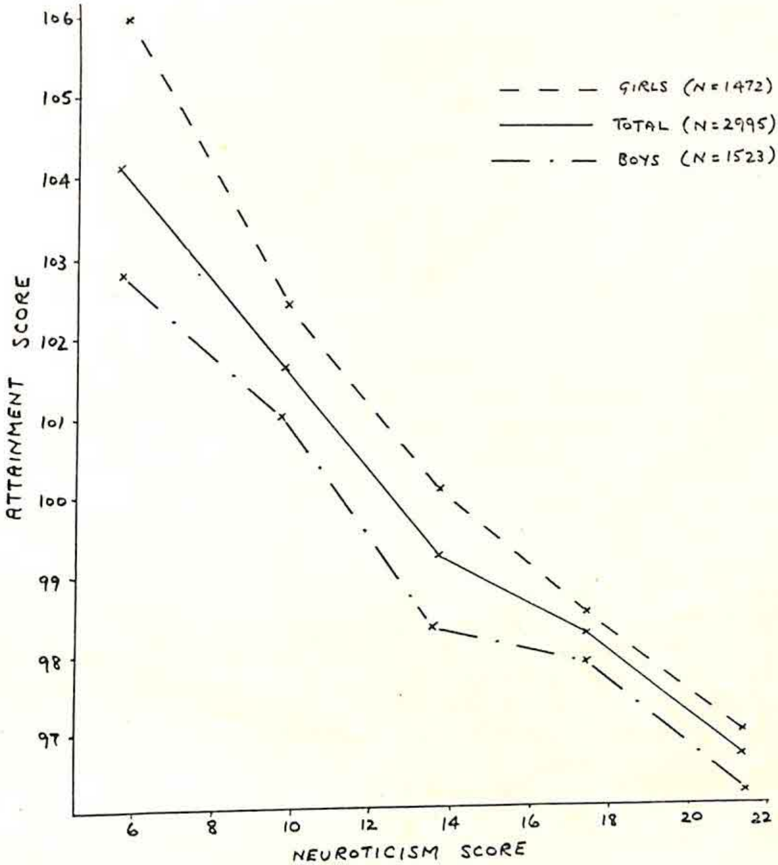


FIGURE 1

REGRESSION LINES FOR ATTAINMENT ON NEUROTICISM.

The same is not true for extraversion, as can be seen from Figure 2. For the whole group there is a U-shaped relationship and this is significantly non-linear (F=4·32 with d.f. of 2990 and 2). This non-linearity explains the low value of the correlation coefficient reported earlier, correlation being inappropriate for describing a non-linear relationship. The analysis by sexes separately explains the shape of the overall curve. It is a composite curve which combines two quite different relationships. For girls there is a positive relationship between extraversion and attainment, but for boys the relationship is negative. Rushton (1966) and Hallworth (1964) found that extraversion and attainment were positively correlated, whereas Child (1964) produced an opposite finding, using a sample which may have included more boys than girls. The sex difference indicated in the above analysis could account for these contradictory findings.

It had become clear by this stage in the investigation that the relationships between the personality dimensions and attainment were far from being simple. To show how attainment varied with extraversion at each level of neuroticism, mean quotients were calculated for nine different personality types for each sex.

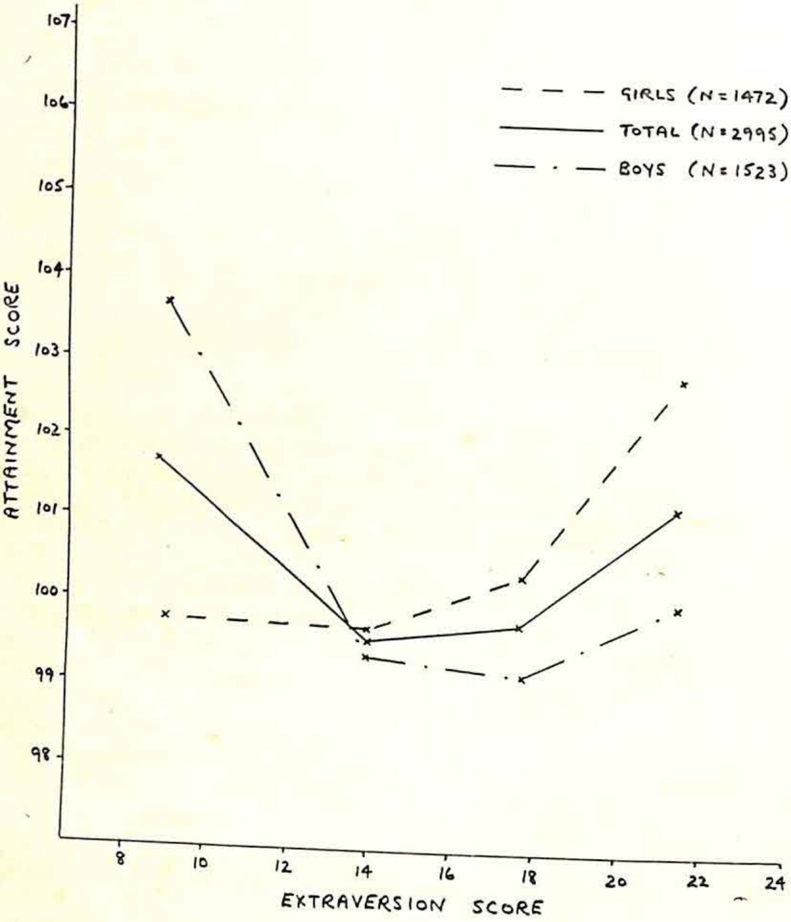


FIGURE 2

REGRESSION LINES FOR ATTAINMENT SCORE ON EXTRAVERSION.

TABLE 5

MEAN SCORES ON SECONDARY SCHOOL CRITERION OF BOYS OF DIFFERING PERSONALITY TYPES.

JEPI Score	High 'E' (24—18)	(N)	Middle 'E' (17—12)	(N)	Low 'E' (11—0)	(N)	All 'E' groups	(N)
High 'N' (24—16)	97·60	(108)	96·80	(73)	102·74	(34)	98·14	(215)
Middle 'N' (15—10)	97·56	(315)	98·14	(206)	97·70	(36)	97·78	(557)
Low 'N' (9—0) ..	101·14	(516)	101·62	(198)	111·06	(37)	101·76	(751)
All 'N' groups ..	99·53	(939)	99·38	(477)	103·92	(107)	99·79	(1523)

TABLE 6

MEAN SCORES ON SECONDARY SCHOOL CRITERION OF GIRLS OF DIFFERING PERSONALITY TYPES.

JEPI Score	High ‘ E ’ (24—18)	(N)	Middle ‘ E ’ (17—12)	(N)	Low ‘ E ’ (11—0)	(N)	All ‘ E ’ groups	(N)
High ‘ N ’ (24—16)	97·28	(110)	98·68	(116)	95·53	(34)	97·68	(260)
Middle ‘ N ’ (15—10)	99·99	(325)	98·72	(273)	100·79	(37)	99·49	(635)
Low ‘ N ’ (9—10) .	104·79	(330)	102·45	(212)	102·57	(35)	103·80	(577)
All ‘ N ’ groups ..	102·67	(765)	100·03	(601)	99·69	(106)	100·86	(1472)

Both Table 5 and Table 6 show that the stable child, whether boy or girl, tends to have a slightly higher position in class than the neurotic child at all levels of extraversion. Table 5 shows that the introverted boy tends to be more successful at all levels of neuroticism, but the boys in the ‘ stable introvert ’ category form a distinctly superior group. Table 6 shows a tendency for the extraverted girl to show a higher level of attainment, but this is less strongly marked. Girls who are ‘ stable extraverts ’ have the highest mean attainment score.

The final stage of this investigation was to test the hypothesis that among the brightest children, the ‘ neurotic introverts ’ would form the superior group on school attainment. For this analysis the punched cards for all those children with attainment quotients of 110 and above were re-sorted into the same personality categories as before. The mean scores for these contrasting personality types among bright boys and girls are shown in Tables 7 and 8.

TABLE 7

MEAN SCORES ON SECONDARY SCHOOL CRITERION OF BRIGHT BOYS OF DIFFERING PERSONALITY TYPES.

JEPI Score	High ‘ E ’ (24—18)	(N)	Middle ‘ E ’ (17—12)	(N)	Low ‘ E ’ (11—0)	(N)	All ‘ E ’ groups	(N)
High ‘ N ’ (24—16)	119·37	(19)	117·10	(11)	118·67	(12)	118·55	(42)
Middle ‘ N ’ (15—10)	116·56	(56)	118·19	(42)	119·92	(12)	117·55	(110)
Low ‘ N ’ (9—0) ..	117·87	(144)	118·49	(67)	123·00	(20)	118·50	(231)
All ‘ N ’ groups ..	117·66	(219)	118·25	(120)	120·98	(44)	118·23	(383)

TABLE 8

MEAN SCORES ON SECONDARY SCHOOL CRITERION OF BRIGHT GIRLS OF DIFFERING PERSONALITY TYPES.

JEPI Scores	High ‘ E ’ (24—18)	(N)	Middle ‘ E ’ (17—12)	(N)	Low ‘ E ’ (11—0)	(N)	All ‘ E ’ groups	(N)
High ‘ N ’ (24—16)	116·53	(21)	118·52	(23)	112·84	(6)	117·00	(50)
Middle ‘ N ’ (15—10)	118·11	(72)	118·29	(70)	118·82	(11)	118·24	(153)
Low ‘ N ’ (9—0) ..	119·17	(120)	119·16	(74)	117·67	(15)	119·06	(209)
All ‘ N ’ groups ..	118·55	(213)	118·71	(167)	117·16	(32)	118·50	(412)

In this final analysis the numbers of children in the cells are, in several instances, too small to allow any significant differences to emerge. In Table 8 the bright girls show exactly the same pattern as the girls in the complete ability range; 'stable extraverts' form the superior group on school attainment. Again it appears that 'stability' of personality is the most important characteristic as far as school attainment is concerned. The extraversion dimension appears to be less closely related to attainment.

Among the bright boys the situation is less clear; the pattern emerging from this analysis is rather different from that obtained with the complete ability range. Among these boys Table 7 shows that stability appears to be less important than introversion. Again the small groups involved demand caution in interpretation, but there is at least a suggestion that introversion is more closely related to attainment in this sample than is stability. It is still true, however, that the group with the highest mean attainment score is that of the 'stable introverts.' There is no suggestion amongst either boys or girls that the 'neurotic introverts' are more successful in their schoolwork.

Tests of statistical significance on the 'all E' and 'all N' groups in Tables 5–8 support the above interpretations. In the complete ability range the differences in mean attainment scores for boys both between the neurotic and stable extreme groups and between the extrovert and introvert extreme groups are significant. In the results for bright boys, however, only the difference between extrovert and introvert is significant. Among the girls in the complete ability range, the difference in mean attainment scores between the neurotic and stable extreme groups is significant, but there is no significant difference on the extraversion dimension. This result was also found with the bright girls.

IV.—CONCLUSIONS AND DISCUSSION.

A sample of over 2,700 children provides results which can be treated with more confidence than those from many of the previous studies with samples which were small and often unrepresentative. The initial four hypotheses oversimplified the pattern of relationships which emerged from this investigation. Simple verification or rejection of each separate hypothesis was impossible, as the following summary of results shows.

(1) Neuroticism shows a significant correlation with school attainment. Children with high neuroticism scores tend to be less successful than those with low neuroticism scores. This is true for both boys and girls. The relationship is not curvilinear and thus the correlation coefficient (—·16) gives a true indication of the degree of relationship.

(2) Extraversion shows no significant correlation with school attainment, because a distinct sex difference produces an overall non-linear relationship. Extraverted girls and introverted boys tend to be more successful in school work than children with the opposite personality characteristics.

(3) Girls who are 'stable extraverts' and boys who are 'stable introverts' show the highest mean attainment scores. This is also true of the brightest children. The hypothesis that 'neurotic introverts' would be the most superior group among bright children must be rejected.

Each of these findings clarifies particular points arising from previous studies. The essential linearity of the relationship between neuroticism and school attainment is contrary to Lynn and Gordon's (1961) findings. They extended the use of the Yerkes-Dodson law to human subjects to predict a non-linear relationship. However, the Yerkes-Dodson law is adaptable; for simple tasks high drive (which Eysenck links with neuroticism) facilitates learning; for tasks of moderate complexity an inverted-U relationship is predicted; while for difficult tasks high drive interferes with learning efficiency. As we do not know which tasks in human learning can be equated with those which are difficult for rodents, no firm predictions can be made from the law. Brown (1965) has suggested that even in animal learning experiments, there is insufficient evidence to substantiate it. The extension into human learning is thus highly speculative. Those who wish to use Eysenck's theoretical framework relating neuroticism to drive level and also the Yerkes-Dodson law relating 'drive level' to task performance will find support from the present findings. But they will have to assume that the teacher confronts his class with 'difficult tasks.' The children at least would support this view.

It has already been suggested that the sex difference on the extraversion dimension may explain previous contradictory findings. But why extraversion should be an advantage for girls in their schoolwork, but a disadvantage for boys, is not at all clear. It may be that girls act out their extraversion, especially stable extraversion, in a more scholastically acceptable way than do the boys. Perhaps the values of the boys' peer groups tend to be in opposition to academic success. Presumably extraverted boys will participate in these groups more actively than the introverts and this might result in a less successful academic record. If the values supported by the girls' peer groups supported academic success, this might account for the sex difference which has been identified.

The superiority of the introverted boy is not difficult to explain in terms of the behaviour patterns ascribed to the introvert by Eysenck (1965a, page 59): "The typical introvert . . . is a quiet retiring sort of person, introspective, fond of books rather than people; he is reserved and distant except with intimate friends. He tends to plan ahead, 'looks before he leaps,' and distrusts the impulse of the moment . . . He keeps his feeling under close control, seldom behaves in an aggressive manner, and does not lose his temper easily. He is reliable, somewhat pessimistic, and places great value on ethical standards."

In terms of Eysenck's theoretical formulation of personality structure, the introvert should also be at an advantage in school learning. Eysenck suggests that the extravert builds up reactive inhibition more quickly than the introvert. This interferes with the learning process and prevents the extravert from being able to concentrate on the learning task for more than short periods of time. Thus, from both behaviour patterns and Eysenck's personality theory, the introvert should reach a higher standard, on average, than the extravert in school work. But why should this be true of boys and not of girls? Does the female introvert withdraw too completely from normal class interaction in the learning process and thus fail to learn as effectively as the introverted boy?

Finally, the hypothesis that, among bright children, 'neurotic introversion' would be associated with scholastic success had to be rejected. The relationships found amongst the university students cannot be simply a result of selecting the brightest school children. Perhaps, as intelligent children grow older, they also become more self-critical and their scores on self-report inventories move towards the neurotic end of the scales. Or perhaps the self-reliance expected at university level actually makes students who take their responsibilities seriously become neurotic.

This study has perhaps thrown some new light on the problem of personality relationships with school attainment, but the impression it must leave is of complex inter-relationships about which we still know all too little.

ACKNOWLEDGEMENTS.—The 'Age of Transfer' project is sponsored by the Scottish Council for Research in Education. We are grateful for the co-operation of the Directors of Education from both Aberdeen town and county. In particular we should like to acknowledge the help of all the many teachers who administered the tests and supplied class orders of merit.

V.—REFERENCES.

BROWN, W. P. (1965). The Yerkes-Dodson law repealed. *Psychol. Reports*, 17, 663-666.

CHILD, D. (1964). The relationships between introversion-extraversion, neuroticism and performance in school examinations. *Brit. J. Educ. Psychol.*, 34, 187-195.

EYSENCK, H. J. (1965a). *Fact and Fiction in Psychology*. London : Penguin Books.

EYSENCK, H. J., and WHITE, P. O. (1964). Personality and the measurement of intelligence. *Brit. J. Educ. Psychol.*, 34, 197-201.

EYSENCK, SYBIL (1965b). *J.E.P.I. Manual*. London : Univ. of London Press.

FURNEAUX, W. D. (1962). The psychologist and the university. *Univ. Quart.*, 17, 33-47.

HALLWORTH, H. J. (1964). Personality ratings of adolescents : a study in a comprehensive school. *Brit. J. Educ. Psychol.*, 34, 171-177.

KELVIN, R., LUCAS, C., and OJHA, A. (1965). The relationship between personality, mental health and academic performance in university students. *Brit. J. Soc. Clin. Psychol.*, 4, 244-253.

LYNN, R., and GORDON, I. (1961). The relationship of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

McNEMAR, Q. (1962). *Psychological Statistics*. New York : Wiley, Toppan.

NISBET, J. D., and ENTWISTLE, N. J. (1966). *The Age of Transfer to Secondary Education*. London : Univ. of London Press.

RUSHTON, J. (1966). The relationship between personality characteristics and scholastic success in 11-year-old children. *Brit. J. Educ. Psychol.*, 36, 178-184.

YATES, A., and PIDGEON, D. A. (1957). *Admission to Grammar Schools*. London: Newnes.

(*Manuscript received 24th May, 1967*)

CREATIVITY AND VERBAL ABILITY: A DIRECT EXAMINATION OF THEIR RELATIONSHIP

BY GERALD P. GINSBURG

AND ROBERT G. WHITTEMORE

(University of Nevada, Reno)

SUMMARY. The threshold concept of the relation between creativity and intelligence implies that measures of creativity begin to differ markedly from measures of IQ only above a certain level of intelligence. This hypothesis was tested by comparing the variance of scores on RAT at different levels of performance on ACT by 292 Nevada students. The hypothesis was not confirmed. A similar finding was got from analysis of the scores of 107 Australian boys.

I.—INTRODUCTION.

IN the twenty-two years of the post-war era, creativity research has burgeoned like a weed. Many of the products of this rapid proliferation have resembled weeds and have been aptly criticized by various psychometric gardeners. One of the central and well deserved areas of criticism is the tendency to reify 'creativity' as an entity independent of intelligence and aptitudes closely related to intelligence (Cureton, 1964; McNemar, 1964; Thorndike, 1963; Wodtke, 1964). It is quite clear that although the problem of separating creativity and intelligence has been raised frequently (Barron, 1963; Getzels and Jackson, 1962; Guilford, 1950; McNemar, 1964; Mednick and Mednick, 1965; Torrance, 1959; Wallach and Kogan, 1964; Yamamoto, 1965, 1966), no adequate resolution has yet occurred. The persistence of the problem is in part due to the wide variety of creativity measures, often poorly validated and with low or unknown reliabilities, and in part to the conceptual confusion and variety subsumed under the rubric 'creativity' (cf. Cureton, 1964; Ginsburg, 1964; Wodtke, 1964). Moreover, as McNemar (1964) has pointed out, the current fad among most creativity investigators is either to discount the role of intelligence or to assume that it is not importantly involved in their performance measures of creativity.

McNemar argued that there should be a triangular scatter diagram between "honest to goodness criterion measures . . . of creativity" and IQ, where having "a high IQ is not a guarantee of being creative; having a low IQ means creativity is impossible" (1964, p. 879). The purpose of this paper is to report a direct application of McNemar's challenge and to discuss the ramifications of the results.

II.—METHOD.

Form I of Mednick's (1962) Remote Associates Test (RAT) was selected as our measure of 'creative potential' (a more accurate term than 'creativity,' for reasons mentioned below), and ACT-Verbal scores were used as a measure of scholastic atptitude. Both tests were administered to 292 entering freshmen at the University of Nevada (Reno) in September, 1965, as part of the entrance examination program. The ACT-Verbal scores are assumed to reflect verbal IQ, which was of direct interest in this study because of the verbal nature of the RAT. This assumption appears justified both by the strong, positive relationship between ACT and IQ scores and by the similar results obtained when ACER Verbal IQ and RAT scores were correlated in another population.

The choice of the Remote Associates Test requires some elaboration to demonstrate the relevance to the problem at hand. The RAT was chosen as a measure of creative potential because it is easy to administer in a group situation, its reliability is well documented and high (corrected odd-even coefficients are about ·91 for Form I; cf. Mednick and Mednick, 1965, p. 16), and its relationship to a wide range of criterion measures, including some artistic criteria, has been investigated carefully (Mednick and Mednick, 1965). Moreover, as pointed out elsewhere (Ginsburg, 1964, Ch. 2), the RAT provides an estimate of how creatively a particular person could perform in a situation which was conducive to creative performance by that person. This interpretation conceives of the RAT as a measure of a person's potential creativeness, independent of situational, task and personal knowledge determinants. However, Wallach and Kogan (1964, p. 7) argue that the RAT, as well as most other measures of creative ability, is administered in a 'test' context which itself is antagonistic to creative performance. Their point is quite reasonable, but it may be more applicable to the identification of situations conducive to creative performance than to the measurement of creative potential. We will return to this point later in the paper.

Earlier work in which the relationships between RAT and various intelligence and scholastic aptitude measures were examined revealed a consistent, moderately strong, positive correlation between RAT scores and verbal subscale scores (cf. Mednick and Mednick, 1965, pp.38-41). Also, Mednick and Mednick (1964) report that the IQ–RAT correlation is much less for IQs above 110 than for the whole obtained range of IQ scores. Obviously, this drop in the size of the correlation could have been an artifact of truncation of the IQ range; but the same implication—that the RAT–IQ correlation breaks down in the higher reaches of the IQ range—was presented by other data in which there was no truncation of range (Ginsburg, 1964). However, Ginsburg's data lent only indirect support to the argument; he had not undertaken a direct examination of the bivariate scatterplots. Yamamoto also provides evidence for what he calls a "threshold of intelligence" (1965, p. 304), where above a certain level of intelligence (a level which may vary from one field to another) the relationship between intelligence and creativity breaks down; but Yamamoto used the Minnesota Tests of Creative Thinking (Torrance, 1962) rather than the RAT.

Thus, the problem is well-focussed: the RAT appears to be a measure of creative potential worthy of serious attention. Moreover, indirect evidence exists, in answer to McNemar's demand (1964), of a triangular shape in the RAT-intelligence bivariate distribution. However, the evidence is very indirect and the triangular shape of the distribution can only be set forth as a hypothesis. So McNemar's challenge still stands, and enough evidence has accumulated to warrant a direct and unambiguous answer. Very simply, we decided to examine scatterplots of RAT-English ACT bivariate distributions to see if there is greater variation among RAT scores at the upper levels of ACT scores than there is at lower levels of ACT scores.

McNemar (personal communication, February, 1965) had recommended that we examine variances within column arrays, but this led to the problem of specifying the criteria by which the column widths should be established. The sizes of the columns could drastically affect the results: equal column widths would yield grossly unequal column *n*s, and equal column *n*s would result in grossly unequal column widths. Departure from homoscedasticity could not be tested by the usual techniques because they would yield no information about the location of the departure(s). However, the problem was resolved through a suggestion from Dr. Warren Norman (University of Michigan, personal communication, May, 1966). He recommended that we compute the RAT deviations about the RAT-on-ACT linear regression line, split the abscissa at a few locations (with approximately equal numbers of cases in each column), and compare the means of the squared deviations within each column. By this technique we could make use of wide columns at the extremes of the ACT distribution (where cases are rare) without taking each deviation in such a wide column from a single average value in the whole array of that column. Instead, every deviation would be computed from its own vertical projection on the regression line. Since visual inspection of the scatterplots revealed no indication of curvilinearity, we followed Norman's suggestion. The results from the University of Nevada sample are presented below; supplementary findings based on other investigators' data will be discussed later.

III.—RESULTS.

The ACT (English) range for the 292 University of Nevada students was 2·1—31·5. The distribution was broken into six columns with *n*s as similar as possible: 2·1—12·3 ($n=43$); 12·9—15·9 ($n=49$); 16·5—18·3 ($n=63$) 18·9—20·1 ($n=43$); 20·7—21·9 ($n=48$); and 22·5—31·5 ($n=46$). The ACT (English) mean was 18·02 and the variance was 22·92. The RAT range was 0–24, mean 10·21, variance 25·58. Table 1 shows the mean of the squared deviations of RAT scores from the linear regression line for each vertical array (product-moment correlation was ·44).

TABLE 1

MEAN SQUARED DEVIATIONS OF RAT SCORES FROM LINEAR REGRESSION LINE WITHIN COLUMNS OF ACT (ENGLISH).

ACT (English) Column Range	n	Average Squared RAT Deviation(a)
(1) 2·1—12·3	43	20·33
(2) 12·9—15·9	49	19·64
(3) 16·5—18·3	63	21·65
(4) 18·9—20·1	43	21·16
(5) 20·7—21·9	48	15·60
(6) 22·5—31·5	46	25·09

(a)—Computed as $(Y—Y')^2/n$, where Y=RAT score.

ACT Col. (6) shows a RAT variation much higher than that of ACT (5). However, the RAT variations in ACT Cols. (1)—(4) are fairly close to that of Col. (6), so that Col. (5) appears to reflect a deviation from the pattern. The large differential between Cols. (5) and (6) is due more to the exceptionally low variation in Col. (5) than to an exceptionally high variation in Col. (6). Therefore, the hypothesis being tested did not receive support. This conclusion will be buttressed by the supplementary data discussed below.

IV.—DISCUSSION.

The data presented above do not support the hypothesis that RAT scores show more scatter in the higher segment of the verbal intelligence range than in the rest of that range. There is some question, however, as to whether the data actually refute the hypothesis, since there is a slight trend in Table 1 of an increasing RAT deviation from lower to higher ACT scores. Consequently, additional data gathered in Australia by Dr. Ronald Taft (University of Melbourne) were subjected to a similar analysis. Taft administered the RAT (Form 1) and the ACER intelligence test, FORM W, to 107 boys aged 15–16 years. We constructed a scatterplot of Taft's data, with the language subscore of the ACER on the abscissa and RAT score on the ordinate, and drew the linear regression line of RAT on ACER. We then broke the ACER distribution into three columns (Col. 1 : range 86–105, $n_1=30$; Col. 2 : range 106–117, $n_2=43$; Col. 3 : range 118–135, $n_3=34$) and measured the average squared RAT deviation around the regression line within each column. The mean squared RAT deviation in Col. 1 was 17·32 ; in Col. 2, 15·66 ; and in Col. 3, 14·81. The product-moment correlation between the two variables was ·60. Taken in conjunction with our University of Nevada data, the hypothesis is refuted. There simply is not a greater variation among RAT scores in the higher segments of the verbal intelligence range than in the remainder of that range.

This conclusion poses an interesting problem in regard to the RAT as a measure of creative ability. The test clearly is positively and linearly related to verbal intelligence measures, and the relationship holds up across the full range of obtained scores on such measures. On the other hand, RAT scores predict well to certain types of performance to which intelligence measures do not predict (Mednick and Mednick, 1965, pp. 159-160). For example, verbal IQ measures did not predict ratings of research creativity of graduate students, whereas RAT did so quite well. The same findings obtained in a validity study of the RAT : despite the positive correlation between the RAT and the Terman Concept Mastery Test, only the RAT predicted instructors' ratings of the creativity of the designs of architecture students. Finally, RAT and verbal intelligence measures were differentially related to college grades, the latter having a moderate positive relationship and the former a low negative relationship.* Therefore, despite the persistent finding of a moderate, positive and reasonably homoscedastic relationship between RAT and verbal intelligence measures, the two appear to be measuring different (albeit overlapping) abilities.

Two other viewpoints which have been presented in print should be considered. Wallach and Kogan (1964) have argued that past studies have provided little justification for talking about creativity and intelligence as separate concepts at the same level of abstraction. They contend, however, that the assessment of creativity in a playful (as opposed to a testing) context would yield high correlations between certain creativity measures and low correlations between the creativity measures and intelligence measures (which by definition impose a testing context on the assessment situation). In fact, their data strongly supported their argument. On the other hand, it is quite possible that the low IQ-creativity correlation which they obtained was due simply to the differences in the assessment context (a testing context *versus* a game context, respectively). Until this possibility is tested directly by some design like the multimethod-multitrait recommendation of Campbell and Fiske (1959), Wallach and Kogan's results must be accepted with due caution. It would be inappropriate at present to accept their argument of a clear separation between creative ability and intelligence, and—in view of the data reported by Mednick and Mednick (1965)—it would be equally inappropriate to accept without question the argument that other studies provide little justification for talking about creativity and intelligence as separate concepts.

---

* The low negative relationship between RAT score and college grades apparently can be explained by the findings that high RAT scorers are more variable in their grades than low RAT scorers (Walker, 1962), and that high RAT scorers obtain better grades from 'flexible' instructors than they do from less 'flexible' instructors, whereas the grades of low RAT scorers are not affected by the 'flexibility' of the instructor (Miller, 1960).

Another current viewpoint about creativity and intelligence is embodied in the 'threshold of intelligence' concept mentioned earlier. MacKinnon's statement, cited by Yamamoto (1965), expresses the idea clearly: "... above a certain required minimum level of intelligence which varies from field to field and in some instances may be surprisingly low, being more intelligent does not guarantee a corresponding increase in creativeness. It just is not true that the more intelligent person is necessarily the more creative one" (MacKinnon, 1964, p. 311). McNemar also implies a threshold concept when he expects creativity measures to demonstrate that "having a high IQ is not a guarantee of being creative; having a low IQ means creativity is impossible" (1964, p. 879). Yamamoto (1965) interprets his data as supporting the threshold concept. However, our data provide no support for the threshold concept, and it would appear worthwhile to examine Yamamoto's findings carefully to ascertain the basis for his inference.

In his Table 1, Yamamoto (1965, p. 302) presents the means, standard deviations and ranges of his IQ and creativity measures for each of the subgroups in his two samples (school systems "A" and "B"). We constructed curves—one for each school system—based on the subgroup means (male and female combined) of IQ and creativity scores, with IQ on the abscissa and creativity on the ordinate. The results are shown in Figure 1.

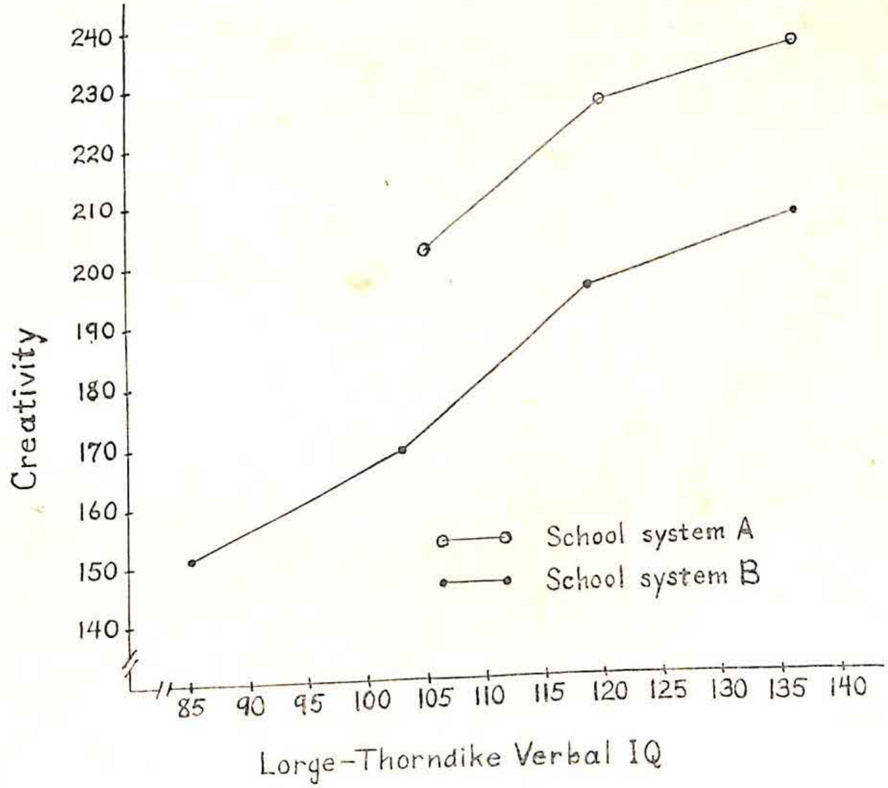


FIGURE 1

CREATIVITY AND VERBAL I.Q. (Adapted from Yamamoto, 1965, p. 302, Table 1).

It can be seen that the curves for the two school systems vary in elevation but apparently not in shape. In each case, the relationship between the two measures is curvilinear, with an inflection point about IQ 120 where the slope of the curve drops dramatically. Furthermore, the variances in creativity scores for the different segments of the IQ range which Yamamoto established do not differ significantly from each other, although the creativity variance does tend to increase slightly as one moves from the lower to the upper segments of the IQ range. Putting together the curvilinearity obtained by a plotting of the IQ and creativity means and the absence of a significantly greater creativity variation within the upper IQ segment relative to the rest of the IQ range, it seems that the threshold concept is *not* supported. That is, the relationship between IQ and creativity does not break down in the upper segments of the IQ range. Instead, the relationship between the two variables is curvilinear, which would explain Yamamoto's finding that " . . . there was a consistent decrease in the size of correlation as the IQ level of subgroups became higher" (1965, p. 304). His findings, then, actually are consistent with the data reported in the present paper, and the threshold concept as defined by a breakdown in the correlation between IQ and creativity cannot be considered supported.

On the other hand, Yamamoto's warning (1965, p. 304) that the case for creative thinking as an entity independent of other facets of intellect should not be overstated is well taken. Creative thinking is a facet of human intellect, but the extent to which it is independent of other facets is still unknown. In our opinion, what is now required is the experimental identification of criterion tasks which relate differently to IQ and creative ability. The validity data reported by Mednick and Mednick (1965), mentioned earlier, are a start in this direction. So are the well-controlled data presented by Mendelsohn and Griswold (1964), where high RAT scorers unwittingly made greater use of incidental stimuli than did low RAT scorers in solving anagram problems. The concept of creative ability requires construct validation, and only by attacking it from different directions can its validity as a construct be assessed. In addition, because of the now established relationship between creative ability and IQ (especially verbal IQ), any study of creative ability must by all means incorporate controls on intelligence.

V.—References.

Barron, F. (1963). *Creativity and Psychological Health*. Princeton : Van Nostrand.

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.*, 56, 81-105.

Cureton, Louise, W. (1964). Creativity research and test theory. *Am. Psychol.*, 19, 136-137.

Getzels, J. W., and Jackson, P. W. (1962). *Creativity and Intelligence*. New York : Wiley.

Ginsburg, G. P. (1964). *Creativity and Childhood Antecedents*. Unpublished doctoral dissertation, Univ. of Michigan.

Guilford, J. P. (1950). Creativity. *Am. Psychol.*, 5, 444-454.

MacKinnon, D. W. (1964). The nature and nurture of creative talent. In R. E. Ripple (Ed.), *Readings in Learning and Human Abilities*. New York : Harper and Row, pp. 305-323.

McNemar, Q. (1964). Lost : Our Intelligence ? Why ? *Am. Psychol.*, 19, 871-882.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychol. Rev.*, 69, 220-232.

Mednick, S. A., and Mednick, Martha, T. (1964). An associationistic view of creative thinking. In C. W. Taylor (Ed.), *Widening Horizons of Creativity*. New York : Wiley, pp. 54-68.

Mednick, S. A., and Mednick, Martha T. (1965). The associative basis of the creative process. Final report, U.S. Office of Education, Co-operative Research Project No. 1073, Ann Arbor : Univ. of Michigan.

MENDELSOHN, G. A., and GRISWOLD, BARBARA B. (1964). Differential use of incidental stimuli in problem solving as a function of creativity. *J. Abnorm. Soc. Psychol.*, 68, 431-436.

MILLER, B. (1960). *A Study of Creativity in College Students and Teaching Method Types.* Senior honours thesis, Univ. of Michigan.

THORNDIKE, R. L. (1963). The measurement of creativity. *Teachers' Coll. Rec.*, 64, 422-424.

TORRANCE, E. P. (1959). Explorations of creative thinking in early school years: a progress report. In C. W. Taylor (Ed.), *The Third* (1959) *University of Utah Research Conference on the Identification of Creative Scientific Talent.* Salt Lake City: Univ. of Utah Press, pp. 58-71.

TORRANCE, E. P. (1962). *Guiding Creative Talent.* Englewood Cliffs, N.J.: Prentice-Hall.

WALKER, H. E. (1962). *Relationships Between Predicted School Behaviour and Measure of Creative Potential.* Unpublished doctoral dissertation, Univ. of Michigan.

WALLACH, M. A., and KOGAN, N. (1964). *A New Look at the Creativity-Intelligence Distinction.* Mimeo.

WODTKE, K. H. (1964). Some data on the reliability and validity of creativity tests at the elementary school level. *Educ. Psychol. Measur.*, 24, 399–408.

YAMAMOTO, K. (1965). "Creativity"—A blind man's report on the elephant. *J. Counsel. Psychol.*, 12, 428-434.

YAMAMOTO, K. (1965). Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. *Brit. J. Educ. Psychol.*, 35, 300-305.

(*Manuscript received* 15*th June*, 1967)

INTELLIGENCE, COMMON ENTRANCE AND THE GENERAL CERTIFICATE OF EDUCATION

By S. W. HOCKEY

(Marlborough College, Wiltshire)

SUMMARY. This report describes an analysis of the school careers of 506 boys between the ages of 13 and 18, those careers being summarised by the marks obtained in the 13+ Common Entrance examination, a Moray House verbal reasoning test, the Ordinary level of the General Certificate of Education, and the Advanced level of the General Certificate of Education. The individual papers of the Common Entrance examination are found, on the whole, to be the best predictors of the corresponding O level subjects, and the IQ appears as the best single predictor of the overall O level performance. The IQ is also by far the best predictor of the overall A level performance. Analysis of the way in which the different school subjects group themselves together at 13 and 16 suggests that there are two significant factors of 13-year-old mental ability, and the factor involving the analytical subjects divides into two by the age of 16.

I.—INTRODUCTION AND PREVIOUS WORK.

THE investigation being reported divides itself naturally into two parts. The first part describes an examination of the predictive value of the Common Entrance examination and of a verbal reasoning test for O and A level GCE, among pupils at Marlborough College. The second part examines, by means of a factor analysis study, the ways in which school subjects group themselves at the different ages of 13 and 16.

There have been numerous studies of the accuracy of 11+ tests for predicting secondary school performance. The Common Entrance (CE) examination has not been subjected to the same scrutiny. The findings of studies such as Stubbins (1940), Peel and Rutter (1951), Emmett and Wilmut (1952), and Wrigley (1955), give an indication of the level of accuracy to be expected in predicting performance in GCE (or School Certificate, as it was formerly called) within a selected group of pupils whose homogeneity reduces correlation.

For example, in these investigations, the correlations between IQ at the age of 11 and performance in specific subjects in GCE or School Certificate at age 16 ranged from +·10 to +·51. When a battery of tests was used in place of the IQ, correlations ranged from +·22 to +·56, and prediction of a criterion of School Certificate performance (combining marks in a selection of subjects) by multiple correlation from a battery of tests gave correlations ranging from +·34 to +·58. The general conclusion of these studies was that IQ is the best single predictor of overall School Certificate (or GCE) performance.

Analysis of the way in which school subjects group themselves together has been carried out less systematically over the years. The earliest workers in the twentieth century had at their disposal only the simplest of statistical tools, and the school subject 'Science' was quite different to that bearing its name today.

But Stead (1923), for example, reported a connection between English, History, Geography and Science (Physics), another between (English) Grammar, Latin and French, and a third in Mathematics (between Arithmetic, Algebra and Geometry).

Later, Ellis (1948) carried out an analysis of nine NUJMB O level subjects, and found evidence for grouping the nine subjects into three principal factors. The (rotated) group factors linked together: (i) Geography, History and English Literature; (ii) Physics, Chemistry and Mathematics, and, rather weakly, Geography; and (iii) (a diffuse grouping) English Language, French and Mathematics.

Later still, Wrigley (1956) presented figures which lead to (rotated) groupings of (i) History, Geography and (General) Science; (ii) English (undifferentiated), Latin, French and History; and (iii) Arithmetic, Algebra, Geometry Science, French and Geography. A similar analysis by Lewis (1961), but with Science divided into Physics, Chemistry and Biology, came to much the same conclusions. The reweighting of Science into three separate subjects led to the resolution of a separate Science and a separate Mathematics component and, at the same time, the remaining subjects (English, Latin, French, History and Geography) clustered together more firmly among themselves. It should also be noted that both Wrigley and Lewis followed a Burt-like approach in analysing correlations only after a factor of general ability had been removed.

The analysis presented below throws rather more light on the relations between the Arts and the Sciences by limiting the weighting of Mathematics and Science to three units out of ten, and by distributing the whole of the correlation variance among the ability factors. The inclusion of an IQ throws light on the psychological nature of the group factors identified, and the addition of the CE battery enables us to study the way in which the features of mental ability change between the ages of 13 and 18.

II.—THE PRESENT INVESTIGATION.

(i) *The examinations.*

The main part of the Common Entrance examination consists of papers set by a central body in Latin, French, English, History, Geography, Scripture and Mathematics. The character of the papers is rather similar to that of the O level GCE papers, and the examination battery is taken by 13-year-old boys wishing to gain admission to one of the public schools. The scripts are marked by the staff of the school to which the candidate concerned is applying for admission.

(ii) *Method.*

The 506 boys whose school careers are analysed in this research entered Marlborough between April, 1958, and January, 1962, inclusive. Their CE papers were marked by the Marlborough staff shortly before their entry, and, shortly after entry, each boy sat the Moray House Verbal Reasoning (Adv.) Test 6. The average age at which the boys sat this test was 13 years 6 months. Subsequently, their GCE O level marks were recorded alongside the figures already available, and finally their GCE A level marks were also recorded.

It will be appreciated that the CE marks were derived from the eleven CE examinations set between February, 1958, and November, 1961, the GCE O level marks from the seven GCE examinations set between July, 1959, and July, 1964, and the A level marks from the five GCE examinations set between July, 1961, and July, 1965. In order to avoid distortion of the results due to fluctuations in the standards of the papers set in these various examinations it was necessary to convert the raw marks to a set of standardised scores. Before this was done, an analysis of the raw score distribution showed that there was insufficient skewness to warrant renormalising.

In the case of the CE examinations the matter is not simple. Common Entrance papers are set three times every year, and the populations which sit the three sets of papers vary quite appreciably in standard. That is to say, if exactly the same papers were taken by the three populations which presented themselves as candidates for admission to Marlborough in any one year, the average scores of those three different populations would differ considerably. In order to standardise the score achieved by any one member of any one population, the assumption was made that the complete set of candidates who present themselves to Marlborough in any one year has, on average, the same ability as the complete set of candidates who present themselves to Marlborough in any other year. With this assumption it is possible to express the average score of any one of the three subpopulations in any one year in terms of the average score obtained by the complete population through the whole year. The simple ratio was then taken as a scaling factor and used to express the average of the sub-population in terms of the whole year's population. A similar method was applied to the standard deviation parameter, and instead of scaling the (CE) raw scores for the subpopulations to means of 10 and to standard deviations of 10, they were scaled to means of 10q and standard deviations of 10s, where q and s are the ratios of the subpopulation means and standard deviations to the corresponding parameters for the whole year's population.

The GCE O level marks were converted directly to suitably ranged z scores, where the z scores were calculated from the equation $z = 10 + 10\ (x - \bar{x})/3\sigma$ the national means ($\bar{x}$) and standard deviations (σ) being supplied by the Oxford and Cambridge Joint Board. The raw A level subject scores were standardised in exactly the same way, and then added for each pupil to give a single figure representing his A level performance. Thus, the A level criterion reflects to some extent the proportions in which the Marlborough pupils opt for the various A level courses, but this was felt to be realistic insomuch as the CE examination is used to some extent as a predictor, even though the final choice of A level course is not known.

Boys who failed to complete their schooling to A level were given an A level score of 0; those who did not take A level for some other reason were excluded from the sample altogether; and those who failed but who completed the course were given a score of 1 for each subject course completed. The standardised scales followed on from here, the maximum score for any pupil in the sample being 18.

The Moray House IQs were not adjusted in any way, and the complete battery of standardised scores was divided into groups in which the combination of O level subjects was the same. This was necessary because Marlborough boys exercise an option on whether they take History or Geography at O level, on whether or not they take Additional Maths as well as Elementary Maths, and on whether or not they take English Literature as well as English Language. The inter-subject and inter-examination correlation coefficients were then computed, and the resulting sets of correlation coefficients were pooled by forming a weighted Fisher's z mean. (No sub-population numbered less than 100.)

III.—RESULTS.

(i) *Prediction of GCE O and A level performance.*

The correlation coefficients in Table 1 were used to calculate the optimum weights of the 13+ scores for predicting GCE O and A level performance. These beta-weights or regression coefficients are set out in Table 2, together with the multiple correlations resulting from the application of these weights.

TABLE 1

CORRELATIONS BETWEEN CE, IQ, GCE O and A LEVEL MARKS.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
CE :																		
1. Latin																		
2. French	49																	
3. English	18	22																
4. History	27	28	21															
5. Geography	17	07	17	32														
6. Scripture	24	19	20	30	22													
7. Maths.	28	32	19	14	09	21												
8. Moray House IQ	21	22	25	16	08	18	47											
GCE O level :																		
9. Latin	53	41	21	17	13	13	38	39										
10. French	32	49	28	21	00	15	26	44	53									
11. Eng. Lang.	23	31	33	24	12	28	20	41	25	42								
12. Eng. Lit.	16	16	26	26	13	21	06	15	24	23	38							
13. History	04	07	09	33	19	13	03	09	20	16	31	35						
14. Geography	10	00	04	28	32	22	19	27	21	15	33	29	48					
15. Elem. Maths.	14	17	17	01	01	08	60	48	33	26	24	10	02	27				
16. Add. Maths.	24	19	13	04	01	05	40	44	41	28	16	03	13	31	72			
17. Science*	08	02	13	05	12	13	30	35	30	25	29	28	20	51	54	62		
18. Number of O level passes	27	28	26	24	13	25	49	45	51	51	52	49	53	61	60	44	60	
19. A level	18	20	24	16	13	11	24	47	47	42	39	33	38	54	34	51	62	59

(Decimal points omitted ; correlations over ·09 are significant at the 5 per cent level.)
* ' Science ' refers to O level Physics-with-Chemistry.

TABLE 2

BETA WEIGHTS FOR O AND A LEVEL CRITERIA AS FUNCTIONS OF 13+ PREDICTORS.

Criterion	Predictor								Multiple R
	CE Latin	French	English	History	Geography	Scripture	Maths.	IQ	
GCE									
Latin	39	14	04	−02	04	−07	13	22	63
French	07	37	12	05	−09	−01	−05	34	61
English Language	01	17	18	07	00	14	−08	32	54
English Literature	04	04	18	16	02	11	−09	09	37
History	−07	01	01	30	10	03	−03	06	35
Geography	00	−13	−10	19	24	10	08	22	47
Elem. Maths.	−02	−01	05	−08	−03	−05	49	27	65
Add. Maths.	14	01	02	−06	−03	−07	23	32	51
Science	01	−11	03	−03	08	04	18	28	41
Number of O level passes	06	05	10	08	02	07	30	23	58
A level	02	07	10	03	07	−03	−01	43	50

(Decimal points omitted.)

(ii) *Factor analysis.*

The pattern underlying the range of O level subjects described in Table 1 is revealed by carrying out a factor analysis. The 17 × 17 (upper lefthand) correlation matrix of Table 1 was subjected to a principal components analysis, and five components with latent roots greater than 1 were found. The five eigenvectors account for 64 per cent of the total variance, and the factor pattern was rotated according to the Promax method. (The Promax rotation gives an oblique solution, and the resulting factors are not necessarily independent.) The results are presented in Table 3.

TABLE 3

PROMAX PATTERN FOR 17 VARIABLES DESCRIBED IN TABLE 1.

GCE	I	II	III	IV	V
Latin	+24	+16	+74	−06	−16
French	+08	+15	+58	−27	+28
English Language	+05	+32	+03	−04	+63
English Literature	−15	+55	+01	−03	+44
History	−13	+82	+07	+08	−05
Geography	+34	+66	−14	+29	−11
Elementary Mathematics	+91	−14	−05	−05	+08
Additional Mathematics	+84	+05	+13	−10	−18
Science	+72	+36	+18	+05	+02
CE					
Latin	−02	−06	+83	+22	−16
French	−10	−10	+80	+02	+09
English	+04	−14	−08	+14	+78
History	−20	+28	+26	+53	+10
Geography	+00	+15	−02	+76	−06
Scripture	+02	−07	−07	+58	+41
Mathematics	+64	−30	+18	+22	+11
Moray House IQ	+56	−06	+07	−01	+36

(Decimal points omitted.)

The rotated solution divides the various subjects neatly into five groups: I (Maths and Science), II (O level History, Geography and English Literature), III (Latin and French), IV (CE Geography, History and Scripture), and V (O level English Language and CE English). We can provisionally label the abilities involved in these groups as: I—abstract reasoning; II—verbal fluency; III—linguistic ability; IV—recall fluency, and V—verbal analysis.

The intercorrelations between the five principal components of Table 3 are such that there is only one second order component with a latent root greater than 1. This component accounts for 38 per cent of the total first order component variance, and correlates most strongly with component III (+0·74) and component V (−0·71).

Analysis of the 13+ correlations leads to only two significant principal components. The Promax solution links the first with (CE) Maths, Moray House IQ, French and Latin, and the second with Geography, History and Scripture. Analysis of the 16+ correlations and Moray House IQ leads to three significant components: the first is associated with (GCE) Maths and Science, the second with History, Geography and English Literature, and the third with French, Latin, Moray House IQ, and English Language.

IV.—DISCUSSION.

We find from Table 1 that the best predictor for any O level subject, with the exceptions of English Language and Additional Maths, is the corresponding CE test. The IQ is a notably successful general predictor, and has particular validity (comparatively speaking) for English Language, Additional Mathematics and Science. It also predicts a comparatively large proportion of the A level criterion variance. As indicated earlier, this latter criterion will be affected by the relative proportions in which the Marlborough population elects to follow the various A level disciplines, but in view of the fact that 50 per cent or so take up A level Mathematics and Science after O level (the rest dividing more or less equally between Languages, the English studies and the Classics), the absence of a significant β-weight between CE Maths and the A level criterion is surprising. The correlation between the IQ and the A level criterion is impressively high considering that the two measures are made some four or five years apart, and that the IQ is essentially a Verbal Reasoning one. The correlation of 0·45 between the IQ and the total number of O level passes gained by each pupil suggests that a candidate's performance at O level depends rather more on the teaching he has experienced and less on his ability than his (subsequent) performance at A level.

Another point of interest arising from the figures of Table 1 concerns the CE English/A level criterion correlation of 0·24. The relative independence of the CE English mark and the IQ leads to a comparatively large CE English β-weight in the A level regression equation (Table 2). We should not perhaps be too surprised that the quality of an A level candidate's work turns out to be a function of his ability in his native language at age 13, whether he eventually becomes a scientist or not.

The nature of the factors underlying the pupils' performance at 13+ and 16+ is fairly clear from the Promax patterns given in Table 3. Factor I can be identified with the non-verbal ability factor commonly accepted by educational psychologists and it is reasonable to suppose that the reasoning component of the Moray House Verbal test is the primary cause behind the moderate loading of the verbal IQ in this factor. We might also conclude that the more numerical items in that test reinforces the association, although the existence of a Latin-French-Maths-IQ factor in the CE battery lends support to the view that numerical reasoning demands an appreciable amount of verbal ability at the age of 13 or so. Presumably the constructors of the Moray House test considered this when compounding their battery of Verbal Reasoning questions. It is also interesting to note the small but significant loading of O level Geography (Table 3, column I), which suggests that Geography becomes more akin to the scientific subjects between CE and O level. We can refer to factor I as one involving *Abstract Reasoning*.

Factor II has high loadings in O level English Literature, History and Geography, and it seems reasonable to label the factor as one involving *Verbal Fluency*. The significantly lower loading of O level English Language can be taken as a reflection of the analytic or grammatical component of that subject, and it is interesting that this factor turns up quite separately from the CE (Geography+History) factor. This could be taken as a further indication that Geography and/or History changes in nature between CE and O level.

Factor III is a very simple factor involving only Latin and French. The nature of these subjects changes little between CE and O level, and we can identify the factor as one involving *Linguistic Ability*. The closeness of the association between the two subjects all the way through the research lends armament to those who feel that the amount of time devoted to the two subjects tends to be disproportionately large in the public schools, and that, in reality, either of the two subjects is capable of giving a sufficient educational experience by itself.

Factor IV is the CE equivalent of the O level Factor II, and the only further point of interest here arises from the fact that at CE, Geography is the leading correlate, whereas at O level, History is the leading correlate. This difference may well be associated with the change in Geography noted above. We can label this CE factor as one involving *Recall Fluency*.

Factor V is essentially connected with O level English Language and CE English, and we can associate it with ability in *Verbal Analysis*. We can also suggest that the ability represented by Factor V is essentially analytical in nature, whereas those represented by Factors II and IV are ones demanding synthesis and possibly a certain amount of factual recall. These contentions are borne out by the separate CE and O level analyses, for at CE, English appears equally associated with the analysis factor of Latin, French and Mathematics, and the synthesis factor of History, Geography and Scripture, and at O level, English Language has equal loadings in the same two factors.

V.—CONCLUSIONS.

(1) In general, the best predictors of O level performance are the corresponding CE papers.

(2) The best single 13+ predictor is the Moray House IQ. Its predictive validity for A level is particularly noteworthy.

(3) The Mathematics test of CE is a somewhat better predictor than the Arithmetic (11+) test used by earlier investigators. The CE English is, however, rather worse than its 11+ English counterparts.

(4) There is evidence for five basic factors of mental ability between 13+ and 16+. One is linked with Mathematics, Science and IQ, a second with GCE English Literature, History, and Geography, a third with Latin and French, a fourth with CE Geography, History and Scripture, and a fifth with CE English and GCE English Language.

ACKNOWLEDGEMENTS.—The author acknowledges with grateful thanks the help which the following have given to this work: Mr. A. E. E. McKenzie, Oxford and Cambridge Schools Examination Board; Mrs. D. M. Penfold, London Institute of Education; Professor F. W. Warburton, Manchester University Department of Education; Dr. H. J. Hallworth, Birmingham University Department of Education; Professor J. Wrigley, Reading University Institute of Education; Mrs. C. E. Taylor, Computation Laboratory, Southampton University; Mr. M. J. Harrison, Marlborough College. The project was commissioned and supported by the Master of Marlborough College, Mr. J. C. Dancy.

VI.—REFERENCES.

ELLIS, A. G. (1948). *The Relations between Abilities in Grammar School Subjects.* M.Ed. Thesis, Durham University.

EMMETT, W. G., and WILMUT, F. S. (1952). The prediction of school certificate performance in specific subjects. *Brit. J. Educ. Psychol.*, 22, 52-62.

LEWIS, D. G. (1961). *An Experimental Study of Ability in Elementary Science.* Ph.D. Thesis, Manchester University.

PEEL, E. A., and RUTTER, D. (1951). The predictive value of the entrance examination as judged by the school certificate examination. *Brit. J. Educ. Psychol.*, 21, 30-35.

STEAD, H. G. (1923). Correlations between school subjects. *Forum of Education*, 1, 27-34.

STUBBINS, T. E. (1940). The prognostic values of the subject of a secondary school entrance examination. *Brit. J. Educ. Psychol.*, 10, 16-24.

WRIGLEY, J. (1955). The relative efficiency of intelligence and attainment test predictors of success in grammar schools. *Brit. J. Educ. Psychol.*, 25, 107-116.

WRIGLEY, J. (1956). *The Factorial Nature of Ability in Elementary Science.* Ph.D. Thesis, Manchester University.

(*Manuscript received* 13*th July*, 1967)

THE STANFORD-BINET, FORM L–M, THIRD REVISION: A LOCAL ENGLISH STUDY OF NORMS, CONCURRENT VALIDITY AND SOCIAL DIFFERENCES

By C. J. PHILLIPS
(*Centre for Child Study, School of Education, University of Birmingham*)

AND W. J. BANNON
(*Senior Educational Psychologist, City of Birmingham*)

SUMMARY. A sample of children at age 11 years in the City of Birmingham was tested on the Stanford-Binet Scale, Form L–M; at the same time a number of other standard tests of ability and educational attainment were given. Results of junior school leaving examinations were made available by the education authority. The sample is compared with the local population for a number of social and educational characteristics. The results indicate that the published norms for the S–B are inappropriate for English children; that the S–B Scale is highly correlated with academic achievement; and that it is greatly influenced by those verbal abilities which distinguish social classes.

I.—INTRODUCTION.

"MANY decisions of vital importance to the welfare of thousands of children are nowadays being based upon the findings of psychologists attached to child guidance clinics. And these findings rest very largely on the presumed accuracy of certain indispensable psychometric tools such as the Binet scale." This was written thirty years ago by P. E. Vernon (1937). He drew attention to the fact that the "ordinary translated version of the Stanford-Binet scale is quite unstandardised in this country." Since that time two revisions, 1937 and 1960, have resulted in many improvements. It is also the case that educational and child psychologists have other tests to use and that they give rather less weight today than they did thirty years ago to the measurement of general intellectual ability in making decisions that affect the educational and social future of the children that come to them for psychological examination. Nonetheless, they need reliable tools to work with and measuring instruments that yield values which mean what they appear to mean.

The lack of such reliable psychometric instruments is still a serious problem in this country. And a serious aspect of this problem arises from the fact that a large proportion of the decisions to which Professor Vernon referred are made not by educational psychologists but by other professionals who use the S–B, whose recommendations do rest very largely on the indices it yields, and who are too little aware of the inaccuracies in the information obtained as to the intellectual status of a child relative to his age peers.

No British study of the S–B, 1960, has appeared so far. The one published review of the research literature on the S–B, Form L–M, Himelstein (1966), includes no reference to this country. The largest study of the 1937 Revision was that conducted by the Scottish Council for Research in Education (1949) of a sample of Scottish children, born in 1936, at age 11 years. Form L of the 1937 Revision was applied to 1,208 children; the mean IQ was 102·5 and the standard deviation 20·1 (Maxwell, 1961). More recently, S. Jones (1962) published the results of testing, on Form L, a sample of eighty London primary school children, with equal numbers of boys and girls ; at age 10 years the mean IQ was 110·2 ,with standard deviation 14·5. The sample was a more homogeneous one than the Scottish survey and the present study ; her means are much higher.

The present study of a sample of Birmingham children is a contribution to knowledge of the parameters of scores in the population of English children at the age of 11 years. It offers also evidence on concurrent validity and on the relationship of the S–B to some social and educational variables. The representativeness of the sample was judged from its characteristics as compared with known sociological and demographic variables of the population, such as sex, social class and family size. Its representative status was assessed also by the performance of the subjects on other measures of cognitive abilities for which norms from reliable sampling of the population are available ; and the results of other tests which the subjects had taken as part of the allocation procedure for secondary education were obtained. This report deals with the Coloured Progressive Matrices, the English Picture Vocabulary Test 2, and Vernon Graded Word Reading, which were individually administered at the same time as Stanford-Binet, Third Revision, Form L–M.

II.—PROCEDURE.

It was decided to identify all children in the City of Birmingham who were born on the first day of December, 1953, and March and June, 1954. Three dates were chosen in the expectation that these would produce the maximum number of children which could be tested. These birth dates ensured that the average age of the subjects would be 11 years 0 months in the last weeks of the school spring term, 1965, when the field work was carried out. The subjects were in the fourth year of junior schooling.

Identification of subjects was facilitated by the City's records of children prepared for the 11+ allocation procedure. In addition, all independent schools received an account of the project and an appeal for co-operation, and the records of all types of special schools were scrutinised. Two small groups of children were not covered ; firstly, children excluded from the educational system ; and secondly, children of Birmingham families who attend private residential schools outside the City.

The search produced 135 subjects ; they will be designated the 'selected sample.' The 'tested sample' of 127 children received the individual psychological examination.

It was beyond the resources of any single organisation in the City to undertake the field work. Psychologists and their students (post-graduate educational and clinical psychologists in training) co-operated, from the City of Birmingham Education Department, the Department of Psychiatry, the Centre for Child Study and the Uffculme Clinic. In all, sixteen testers took part ; analyses of variance for difference between testers and groups of testers revealed no significant differences for all variables. Most of the examinations were carried out in the subjects' schools in school hours ; but a few were done in the child's home or in a clinic.

The head teachers of the schools were asked to supply the following information concerning each child :

(*a*) The child's position in family, including step-children and adopted children.

(*b*) Father's occupation. They were asked to state if he was disabled or dead, and if there was no father in the home to give other information on the occupational-social-class level of the family.

(*c*) Native English speaking. They were asked to give details if the family was not English-speaking and if the child was not born in this country.

(*d*) Age of entry into Infant School. They were asked for details if the child had not attended an English school continuously from the age of 5 years.

The school was asked also for information concerning disabilities such as poor vision and hearing, and emotional maladjustment ; this information was made known to the psychologist before he examined the child.

The testing procedure was finalised after a number of rehearsals in schools with subjects whose birthdates did not bring them into the survey. The testing session began with the Drawing-of-a-Man test which was followed by the S–B ; testers were recommended to apply the other tests in a particular sequence but were free to vary the order if rapport or other circumstances made it desirable. In the administration of the S–B the following procedures were laid down. Testing began with 'Vocabulary' ; this subtest was continued beyond six consecutive failures at the examiner's discretion, to allow for probable misplacements in order of difficulty for English children. The total vocabulary score was used to determine the point of entry in the Scale ; tests were then presented in the sequential order in which they appear in the Manual until a 'ceiling' level was established, after which continued testing established the 'basal' level. In the case of a success above an already established 'ceiling' or a failure below an already established 'basal' level, testing continued until new levels were obtained. The alternative item at the end of each test year was administered but the results of these subtests were not taken into account in determining the subject's 'basal' and 'ceiling' levels, and they were not used in calculating the mental-ages and IQs used in this report.

In the Vernon Graded Word Reading test the subject was asked to read the words in the left-hand column ; normal testing, by reading the words line by line across the page, began two lines above the first error and continued until all words in two consecutive lines were failed, after which further reading of easier words established a 'basal' level, defined as two consecutive lines read correctly.

Some modifications were made to the instructions in the Manual of Coloured Progressive Matrices to suit individual administration and the child recorded his own answers after the first four items.

The Birmingham School Psychological Service searched files to ensure that no subject had been tested within the previous six months, and testers were asked to report any case in which comments made by the child or his teachers suggested that he was familiar with any item in the testing programme.

III.—CHARACTERISTICS OF THE SAMPLE.

Some characteristics of this sample of 11-year-old children are reported in order to demonstrate its representativeness of Birmingham school children, and to make it possible to compare the results of this survey with others. Information concerning the local population was obtained from the City of Birmingham Abstract of Statistics, which is cited hereafter as 'Birmingham, 1963.'

Of the selected group of 135 children, eight were not tested. These 'lost' subjects were :

(*a*) Three children who were not available ; one was a chronic absentee from school ; one came from a family that had moved from the City ; and in one case parents refused co-operation.

(*b*) Five children who were recent immigrants to Britain, and who had insufficient English language to be tested.

Of the 127 children tested, five were immigrants who had attended English schools for at least four years and had sufficient English language for testing, the remainder were native born.

The expected size of the sample was estimated from the total number of children aged 9 and 10 years at January, 1964, on the registers of primary and special schools (Birmingham, 1963), namely 15,090 at age 9 and 14,999 at age 10. The numbers of children at independent schools in Birmingham are not available but are estimated at 3 per cent of the age group 5 to 14 (Birmingham, 1963, p. 141). Using this estimate, the expected number in the sample was 128, including five from independent schools. The tested sample number 127, including four from independent schools.

The distribution of the sample by date of birth, by type of school, by sex and by father's occupation is set out in Table 1.; the selected and the tested samples may be compared with characteristics of the local population. There is a relative over-representation of births in March and under-representation of those in June. The proportionate numbers of boys and girls attending independent schools are not known; it has been assumed that the ratio of the sexes is the same as in LEA schools (Birmingham, 1963, Tables 100 and 101). Girls are somewhat over-represented in this study. The occupations of the fathers have been classified by the five social classes of the Registrar General, 1961 Census; the information available was too imprecise for reliable classification in some instances. In the absence of information on the distribution of occupations among the fathers of school children in Birmingham, the numbers of live births to Birmingham residents in 1964 classified by social class were used as a basis of comparison. The figures in Table 1 suggest a slight bias in favour of the more highly skilled and socially privileged; the children of unskilled workers may be under-represented. Three pairs of twins were tested; this is approximately the expected number.

TABLE 1

PERCENTAGE DISTRIBUTION OF SAMPLE BY DATE OF BIRTH, TYPE OF SCHOOL, SEX AND FATHER'S OCCUPATION (REGISTRAR-GENERAL'S CLASSIFICATION BY SOCIAL CLASS).

	Expected	Selected Sample	Tested Sample
	%	%	%
Date of birth:			
1st December, 1953	30·5	31·1	30·7
1st March, 1954	35·5	38·5	38·6
1st June, 1954	34·0	30·4	30·7
Type of school:			
County Primary	72·4	73·3	71·7
Voluntary Primary	22·8	21·5	22·8
Special	1·8	2·2	2·4
Independent	3·0	2·9	3·1
Sex:			
Boys	51·3	48·9	48·0
Girls	48·7	51·1	52·0
Father's Occupation:			
I Professional and Managerial	2·0	0·7	0·8
II Skilled (Salaried)	6·3	10·4	11·0
III Skilled (Unsalaried)	60·1	57·8	59·8
IV Partly skilled	13·2	13·3	14·2
V Unskilled	12·2	9·6	8·7
Unknown	6·1	8·1	5·5

Of the 127 children in our tested sample three did not take the junior school leaving examination. One of these, a boy of very superior intelligence, was attending a direct grant grammar school; the other two children were in special schools. The City of Birmingham's allocation procedure for secondary school placement included four tests which were taken a few weeks before this survey. Two tests, Verbal Reasoning 1 and 2, are the Moray House Tests 79 and 80; the tests of Mathematics and English are local products. The scores on the four tests are subject to age allowances covering a twelve months span, and are standardised for the local population to a mean of 100 and a standard deviation of 15. The mean standardised scores and standard deviations of the sample are shown in Table 2.

TABLE 2

MEAN SCORES IN BIRMINGHAM JUNIOR SCHOOL LEAVING EXAMINATIONS.

	Boys (N=59)		Girls (N=65)		Sexes Combined	
	M	s.d.	M	s.d.	M	s.d.
Verbal Reasoning 1 (M.H.T. 79)	100·44	18·20	98·25	14·72	99·29	16·50
Verbal Reasoning 2 (M.H.T. 80)	101·03	17·51	98·92	14·88	99·93	16·22
Mathematics	102·69	15·66	96·17	13·29	99·27	14·83
English	100·64	14·99	100·01	13·88	100·31	14·42

It is possible to make a comparison between local and national performance in the case of Verbal Reasoning Test 1 (MHT 79) from figures kindly supplied by the Godfrey Thomson Unit for Educational Research. The comparison is presented in Table 3.

TABLE 3

MEAN SCORES IN VERBAL REASONING TEST 1 (M.H.T. 79).

	Birmingham norms		National norms	
	M	s.d.	M	s.d.
National sample	102·3	14·7	100·0	15·0
Birmingham sample	100·0	15·0	97·7	15·3
This study sample	99·3	16·5		

There is clear evidence from these comparisons that the local population which has been sampled in this study is unrepresentative of the national population on a group of verbal reasoning; it is probable that the mean of the local sample on S–B also is below the national mean.

The results for boys and girls separately are shown in Table 2. The mean score for boys on the Mathematics Test is superior to that for girls, approaching 1 per cent level of significance; the difference is in the expected direction but it exceeds the expected magnitude. Unexpected is the apparent inferiority of girls on VR 1 and 2; the Godfrey Thomson Unit for Educational Research reports that on average girls have done significantly better than boys on these tests, both in the local and national contexts (personal communication). This is the only evidence we have that casts doubt on the goodness of our sample. There is a chance that the mean scores on all tests taken by girls in the sample are less than the local population means; the matter is not redressed in the 'Reduced Sample', to be explained below, in which boys are still marginally superior, but not significantly. The main implication is an imbalance of abilities between the sexes in our sample which may extend to all measures and affect all comparisons between the performances of boys and girls.

IV.—RESULTS.

(1) *The Stanford-Binet Scale*, 1960 *Revision, Form L-M.*

TABLE 4

TEST RESULTS FOR COMPLETE AND REDUCED SAMPLES.

	Tested Sample (N=127)		Reduced Sample (N=104)	
	Means	s.d.	Means	s.d.
Stanford-Binet:				
Mental-age in months	141·09	29·40	144·20	29·38
I.Q.	104·02	20·02	106·17	19·95
Vocabulary score	14·46	4·62	15·00	4·52
English Picture Vocabulary Test 2: Standardised score	92·95	18·17	94·89	17·79
Vernon Graded Word Reading: Raw score	59·39	25·06	62·74	22·50
Progressive Matrices: Raw score	28·04	5·84	28·56	5·64
Age in months	132·44	2·35	132·38	2·41

Mean scores from the testing of the Birmingham sample are set out in Table 4. (The composition of the 'Reduced Sample' is explained below.) There is a consistent trend of evidence that the mean IQ of English children on the S–B is higher than that of American children and, therefore, that the norms are misleading when used in this country. Equally important is the probability that the standard deviation is significantly greater than the expected figure of 16 points, the basis for the Pinneau deviation IQs in the 1960 Revision. The subjects in this study seem to be a satisfactory sample of Birmingham children at age 11 years. If they are representative of English children in general considerable rescaling of IQs would be needed in use of the 1960 Revision, as was the case in the earlier American versions of the test. For example, the expectation that 16 per cent of children have IQs below 84 would hold, but the top 16 per cent of the population would have IQs above 124, not above 116. The classification of IQs, following the model of Table 1 of the Manual (Terman and Merrill, 1960, page 34), adapated to common English usage, and preserving the same percentage frequencies, would read approximately as follows :

Very superior	154+
Superior	129—153
High Average	117—128
Average	92—116
Low Average	79—91
Dull and backward	66—78
ESN (2·4 per cent)	41—65
Mentally subnormal (0·23 per cent)	40—

It is, of course, not suggested that these categories should be adopted on the basis of this study alone.

The matter may be further confounded by negative skewness in the distribution of IQs. In this sample the distribution departs from symmetry very significantly ($t=4{\cdot}85$; $p<0{\cdot}1$ per cent) ; in the range ½ to 1½ standard deviations there were more cases below the mean (28 per cent) and fewer cases in the corresponding range above the mean (20 per cent), than the expected frequency (24 per cent).

(2) *Stanford-Binet Vocabulary Sub-test.*

The number of words which a child is able to give evidence of understanding is of particular interest as a measure of intellectual development because of the high correlations which are reported with almost all other types of ability testing. Our correlation between the raw score on the vocabulary test and the total raw score, that is mental-age, is ·91, which is very high, but spuriously high because the former measure is contained in the latter. Vocabulary tests, including that in the S–B, 1937 Revision, were most fully studied in England in recent years by Dunsdon and Roberts (1953, 1955, 1957). A sample of 450 children in the age-range of 7 to 15 years in Bristol received the S–B ; they found a correlation with vocabulary score of ·86 and ·83 at ages 10 and 11 years. The mean of the total tested sample was 14·5 words which corresponds very closely to that for American children (see Table 4).

(3) *Vernon Graded Word Reading Test.*

Tests of word recognition are frequently used by teachers and psychologists; the test by Vernon (1938) was selected for this survey because it was less likely to be familiar to the subjects than the similar tests by Burt and Schonell. All these tests have high reliabilities and correlate highly with other measures of reading skill.

The mean score of 59·4 words is almost exactly the expected score at 11 years. Although the test was standardised many years ago it appears to continue to reflect current reading attainments in the age-range with which we are concerned. A notable feature of the Birmingham study is the very large standard deviation, which expressed in terms of ' reading-age ' is 2½ years.

It may be argued that a word recognition test is not an adequate measure of all important aspects of reading ability. We have no evidence to offer on this issue other than the correlation coefficients shown in Table 5, the most impressive of these being the correlation of ·9 with the English Test in the secondary allocation examinations. The correlation of word recognition with S–B mental-age is ·79, which is of the same order as most other investigations that have reported on the relationship between these two tests.

(4) *The Coloured Progressive Matrices.*

The means and standard deviations for this test are given in Table 4. The range of scores was from 13 to 36, the maximum. The distribution of scores in the Birmingham sample shows that the abbreviated version of the Matrices test (Sets A, Ab and B) differentiates very well in the lower ranges of ability It cannot do the same among children of better than average ability at age 11 years: 20 per cent of the sample made not more than two errors. The Manual (Raven, 1958, Table IX) gives norms foı children " over 5 and under 11½ years of age "; the score at the 50th percentile for 11-year-olds is 28. In this study it was 29. The results of the present study generally confirm those norms, but suggest that some modification may be needed at the lower end. The Manual reports a correlation of ·66 with S–B, Form L, for a sample of children at age 9 years. In this study, a very similar result was obtained at age 11 years; the correlation with S–B, Form L–M was ·68. The relationship is somewhat depressed by the ' ceiling ' effect referred to above.

(5) *English Picture Vocabulary Test* 2.

This test, an English adaptation of the Peabody Picture Vocabulary Test, is the only item included in the individual testing programme with norms obtained from English children in recent times. Some importance was attached to its inclusion as it might provide evidence of the representativeness of the sample beyond the local population. Brimer and Dunn (1963) report results for two age groups, 10 : 0—10 : 11 and 11 : 0—11 : 11, based upon the testing of more than 2,000 children. The approximate mean raw score from these results is 27 at age 11 years. The Birmingham sample had a mean of only 22·3. We have no explanation for this finding. Lower mean raw scores were characteristic of the sexes separately: for boys 25·0 as compared with 28·0, and for girls 19·7 as compared with 25·3. The results are shown in Table 4; the mean standardised score of 93 is very much below the expectation of 100, and the standard deviation of 18·2 is significantly greater than 15.

Product-moment correlation of the EPVT with S–B at ·80, is the highest of all the coefficients among the individually administered tests in this study supporting the view that the EPVT is one of the most useful short measures of verbal ability available in this country.

(6) *The relationship between the Stanford-Binet and other tests.*

Some of the tests in this survey have norms expressed in standardised scores for each month of age, other tests such as the Graded Word Reading and Progressive Matrices do not have such norms. Therefore, correlation coefficients have been based upon raw scores. In all the tests there were small positive correlations with age; partial correlations were calculated to eliminate the effect of age differences; this procedure made very small changes to the original correlation matrix. The results are given in Table 5. The correlations between S–B and school achievement tests are very high; the coefficients of ·85 with the two secondary school allocation examinations, Mathematics and English, which are our principal measures of academic attainment, are much higher than the results reported by Himelstein (1966) and as high as the correlations with the two verbal reasoning tests. Indeed, these coefficients are probably not far below the reliability of the S–B. The main conclusion to be drawn from Table 5 is that the S–B, Form L–M, is a very good predictor of academic school learning.

TABLE 5

PARTIAL CORRELATION MATRIX : AGE REMOVED
(N=124) (decimal points omitted)

	Stanford-Binet : Mental-age	Stanford-Binet : Vocabulary	EPVT 2 Raw Score	Vernon G. W. Reading	Progressive Matrices	V.R.I.	V.R. II	Maths.
Stanford-Binet : Vocabulary ..	902							
E.P.V.T. 2 : Raw Score ..	805	820						
Vernon, G. W. : Reading	768	772	701					
Progressive Matrices	678	581	642	508				
V.R.I	852	785	713	841	631			
V.R.II	860	797	720	868	624	968		
Mathematics ...	853	754	713	809	642	914	906	
English	848	832	739	898	595	902	916	865

(7) *Social Class Differences.*

Analysis of test scores for children of different social classes has been made solely in terms of paternal occupation. The socio-economic groups of the 1960 Census were used and combined to form four categories, to conform to the practice in Birmingham (1963) :

I. Employers and managers of large and small establishments ; professional, self-employed and employees ; intermediate and junior non-manual workers. (SEG 1 to 6).
II. Skilled manual ; foremen and supervisors, manual ; own account other than professional. (SEG 8, 9 and 12.)
III. Semi-skilled manual ; personal service workers. (SEG 7 and 10.)
IV. Unskilled manual and unknown. (SEG 11 and 17.)

Farmers, agricultural workers and armed forces (SEG 13 to 16) were not represented in the sample. The distribution of children from families categorised in this way is shown in Table 6.

TABLE 6

SAMPLE : DISTRIBUTION BY SOCIO-ECONOMIC GROUPS.

	Occupied males in Birmingham (1961)	Children : Tested Sample	Children : Reduced Sample
	%	%	%
I	24	17	18
II	46	53	53
III	19	16	14
IV	11	14	15

The first column shows the percentage of occupied males in Birmingham in these four categories (Birmingham, 1964, Table 56), including a small representation of socioeconomic groups 13—16. This information affords one a rough evaluation of the representativeness of this sample of the local child population. It must be borne in mind that children have been sampled, not the offspring of groups of economically active men.

In Table 7 are set out the means and standard deviations of the four socioeconomic groups on all tests, and the F-ratios, all of which exceed the values for a probability of ·001. Without exception the differences between the means are highly significant. The extent of the differences between the offspring of socioeconomic groups is more vividly appreciated from the range of S–B IQs. No child from Group IV achieved the mean IQ of the sample ; the range was 59 to 101. Only two children from Group I failed to achieve the mean of the sample ; the range was from 81 to 156.

TABLE 7

TEST RESULTS : SOCIO-ECONOMIC GROUPS.

	I		II		III		IV		
	Groups 1–6 (N=22)		Groups 8, 9 and 12 (N=67)		Groups 7 and 10 (N=20)		Groups 11 and 17 (N=18)		
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	F =
Stanford-Binet :									
I.Q.	123·27	18·74	104·27	18·42	98·20	14·20	86·06	9·86	16·57
Vocabulary	18·59	4·55	14·46	4·27	13·55	3·68	10·44	2·11	13·80
English Picture Vocabulary Test 2 :									
Standardised Score	108·32	15·45	93·72	16·30	87·30	17·51	77·61	12·12	13·02
Vernon Graded Word Reading :									
Raw Score	81·18	21·66	57·40	24·57	55·60	20·30	44·39	17·78	9·54
Progressive Matrices :									
Raw Score	32·32	3·78	28·36	5·66	24·25	6·02	25·83	4·42	9·08
*Selection Examinations (with age allowance) :									
V.R. I	115·95	10·77	99·20	15·65	95·55	13·10	85·28	11·74	15·35
V.R. II	115·50	11·57	100·56	15·16	95·25	13·48	85·50	10·78	15·38
Mathematics	112·80	12·47	100·09	14·42	93·55	11·11	87·61	8·06	13·31
English	115·90	10·55	99·91	13·75	96·30	11·44	88·94	7·21	16·46

* These tests are based on **N**=124 ; three subjects did not take the tests.

With such great differences between the performance of children from different social backgrounds, one is impelled to go further and to enquire if comparison between each social class on each test yields differences which are beyond the probability of chance. The significance between all pairs of means has been tested by the Newman-Keuls method (Winer, 1962).

The superiority of children from Group I on all tests accords with general expectations; it may be the outcome of the highly selective pressures in our society making for their superior endowment of ability, and of the considerable advantages for intellectual growth in their homes and often in their schools as well. But what of the other three groups who make up more than four-fifths of the population? Although there is evidently a social class variable running through all the tests in Table 7, it does not affect all tests to the same degree. Tests of intellectual attainment with a high verbal content such as Verbal Reasoning, English and the vocabulary tests discriminate between all social groups and the least privileged group is greatly inferior. It is notable that the S–B is such a test. The Progressive Matrices and the Mathematics tests discriminate far less sharply between the three lower social groups, and the test of reading skill does not distinguish them significantly. These results give some support to the argument that it is the development of linguistic abilities and verbal concepts which characterises most sharply the children of different socio-economic backgrounds. The S–B test, despite the fact of its individual administration, does not overcome the disadvantage of this social bias: it reflects verbal abilities throughout the whole social class spectrum.

(8) *The Reduced Sample.*

Samples may be taken from the population of all children in the community at specified ages, selected by methods believed to be independent of the parameters, and otherwise not limited by any other consideration than practicality of testing. On the other hand, there is a case for knowing the parameters of an 'ideal' population from which certain conditions have been specifically excluded. We offer two sets of results. The first comes from the 'Tested Sample' described above. The second is a reduced sample (N=104), which is the result of the following exclusions:

(*a*) Four children described as seriously maladjusted and who were attending a child guidance clinic or the children's psychiatric department of a hospital.

(*b*) Four children who were attending or had attended a special school.

(*c*) Seven children who were immigrants. Three were born in Jamaica and one of these did not enter an English school until the age of 8. One began schooling in British Guiana. Three were children of Indian parentage—in two cases very little English was spoken at home. Not all these children were of low IQ, the range was 69 to 130.

(*d*) Six children who were twins. We do not know for a fact that all the others were single birth, only that they have no surviving twin.

(*e*) We eliminated also those children who were reported as having some serious health problem or physical disability, such as the child with a bad heart condition who had spent many years in hospital, and the girl with severe cataract. But all these cases occurred in one or other of the four classes above and therefore resulted in no further reduction of numbers.

These exclusions left fifty-two boys and fifty-four girls. In order to equalise the representation of the sexes, two more girls were discarded; they were chosen randomly from girls born in March and whose fathers were skilled manual workers, since these two categories were over-represented.

The results of the tests for the Tested Sample and the Reduced Sample are given in Table 4. The mean age for both samples is 11 years 0 months and the range is three months below and above. The Reduced Sample is believed to represent the population of Birmingham children who have attended normal school for six years, who were born in this country and whose native language is English, who are not seriously disturbed emotionally or physically handicapped, and who are not from multiple births. The means for the Reduced Sample are significantly higher than those for the Tested Sample in all the tests. Thus, regional differences apart, when one balances the sexes and excludes certain classes of subjects who might be excluded in the standardisation of a mental test, the estimated mean S–B IQ for English children at 11 years is 106 and the standard deviation is 20 points.

(9) *Secondary Schooling.*

When these subjects were tested individually, they were in their last year at junior school; in the following year the ecucation authority supplied information on the type of secondary schooling which each subject was pursuing. Thirty-seven children were in grammar and other types of selective secondary schools. This group formed 29 per cent. of the sample; the corresponding percentage for the Birmingham population was 27·7.

The official secondary allocation examinations discriminate significantly the ‘ grammar ’ school subjects from those in other types of ‘ selective ’ schools, but the mean S–B IQs of the two groups are not significantly different.

Our principal interest, however, is the comparison between the combined ‘ selective ’ and the ‘ unselective ’ secondary destinations of our subjects.

TABLE 8

TEST RESULTS : SELECTIVE AND UNSELECTIVE SECONDARY EDUCATION.

	Selective Schools (N=37)		Unselective Schools (N=90)		r_{bis}
	M	s.d.	M	s.d.	
S–B I.Q.	126·7	13·0	94·7	14·0	·963
EPVT S.S.	111·5	11·6	85·3	14·5	·869
Vernon, G.W. Reading	83·5	13·3	49·5	21·8	·816
Progressive Matrices	33·0	5·6	26·0	2·4	·718
V.R.I	117·8	7·3	91·7	12·8	·950
V.R.II	118·1	8·0	92·5	12·4	·949
Mathematics	116·3	8·7	92·3	10·6	·972
English	116·7	7·8	93·6	10·6	·964

Bi-serial correlation coefficients are shown in Table 8 together with means and standard deviations of test results for the sample classified by type of secondary school destination. As one would expect, the bi-serial coefficients for the allocation examinations are all very high, ranging from ·95 to ·97. These results provide a basis for evaluating the other tests in the survey. The bi-serial coefficient of the S–B with type of secondary school, ·96, is of the same high order as the allocation examinations; it is much the highest of the bi-serial values for the individually administered tests. These results provide further evidence that the S–B at age 11 years is very highly related to achievement in academic education. All children with IQs of 123 and over entered selective secondary schools; the lowest IQ among these subjects was 105. The cut-off which produced the smallest proportion of misplacements was an IQ of 120.

V.—DISCUSSION AND SUMMARY.

One may be reasonably confident that the mean and standard deviation of the S–B intelligence quotients obtained are good estimates of the local population parameters. Evidence within this study and from independent sources indicates that the mean IQ of the national populatisn may be higher than this local estimate. The standard deviation of IQs among English children in general may be as high as that obtained in this survey. If these conjectures were confirmed, it would follow that use of the American norms of the S–B, Form L–M, must result in serious misclassification of children with respect to their relative intellectual status, particularly in the psychological assessment of children above and below the average range of ability.

Reliable English norms are needed not only for clinical decisions but also for epidemiological studies and other research. For example, McDonald (1967) used the S–B in a study of English children of birth weight 4-lb. or less. She decided that " the sample on which the Terman and Merrill revision was standardised in the United States, in 1937, appears to be the most satisfactory one with which to compare the distribution of IQs " among her subjects. In the absence of better information this decision may have been the only possibility, but its disadvantages are many and manifest.

At the top of the junior school the test is very highly correlated with academic performance, that is in those achievements which are associated with grammar school education. E. L. Thorndike (1921) wrote that for a child who has normal educational opportunities the Stanford-Binet " will prophesy rather accurately how well he will respond to intellectual demands in cases of ' book-learning,' at the time and for some time thereafter, and very possibly for all his life. It will prophesy less accurately how well he will respond to thinking about a machine that he tends, crops that he grows, merchandise that he sells, and other concrete realities that he encounters in the laboratory, field, shop and office. It may prophesy still less accurately how well he will succeed in thinking about people and their passions and in responding to these." The trend in successive revisions of the S–B, based partly on internal validity, has been towards verbal-symbolic content. The results reported above show a decided academic bias which suggests that Thorndike's criticisms of an earlier version are pertinent to Form L–M.

Social class differences in S–B IQs are very large. The view that cultural experience plays a considerable part in accounting for these differences is supported by the finding that in common with other tests of high verbal and linguistic content the S–B reflects the whole range of the social class continuum, to a greater extent than tests that are less verbal. A Working Party set up by the British Psychological Society in 1954 (B.P.S., 1958), included in its report "A Critical Appraisal " of the 1937 Revision of the Stanford-Binet. Among the points of criticism which could still apply to the 1960 Revision is the following : " It markedly penalises the dull person from poor or adverse surroundings, and although this indicates quite well his verbal handicaps it does not give anything like an all-round assessment of intellectual capacity."

All these points are not so much adverse criticism of the nature and content of the test, which rightly retains its place in the repertoire of the educational and child psychologist, as an indictment of current usage. Its norms do not apply to the population of British children ; its concurrent and predictive validity is restricted to a narrower range of learning materials and situations than those for which it is often used. Two qualifications to the above remarks are needed. This study is based upon the result of examining children at 11 years of age and the findings may be generalised to some reasonable range of age but, of course, not down to the pre-school years. The very high relationship of the S–B with academic achievement should not be thought to apply to the acquisition of the basic skills of literacy and numeracy with which S–B may have much less relevance.

ACKNOWLEDGEMENT.—We acknowledge with gratitude the use of computer programmes for the calculation of means, standard deviations and correlations by Dr. H. J. Hallworth, School of Education, and of analyses of variance by Dr. P. M. Levy, Department of Psychology, The University, Birmingham.

VI.—REFERENCES.

BIRMINGHAM, CITY OF (1963). *Abstract of Statistics: Number Eight.* (Editor, Neale, A. B.). City of Birmingham Central Statistical Office.

BIRMINGHAM, CITY OF (1964). *Abstract of Statistics: Number Nine.* (Editor, Neale, A. B.). City of Birmingham Central Statistical Office.

BRIMER, M. A., and DUNN, L. M. (1963). *Manual for the English Picture Vocabulary Tests.* London: N.F.E.R.

BRITISH PSYCHOLOGICAL SOCIETY (1958). Appendix I—the 1937 version of the Stanford-Binet Scale—a critical appraisal. *Bull. Brit. Psychol. Soc.*, 35, 13-15.

DUNSDON, M. I., and ROBERTS, J. A. FRASER (1953). The relation of the Terman-Merrill Vocabulary Test to mental age in a sample of English children. *Brit. J. Statist. Psychol.*, 6, 61-70.

DUNSDON, M. I., and ROBERTS, J. A. FRASER (1955). A study of the performance of 2,000 chiildren on four vocabulary tests; I.—Growth curves and sex differences. *Brit. J. Statist. Psychol.*, 8, 3-15.

DUNSDON, M. I., and ROBERTS, J. A. FRASER (1957). A study of the performance of 2,000 children on four vocabulary tests; II.—Norms, with some observations of the relative variability of boys and girls. *Brit. J. Statist. Psychol.*, 10, 1-16.

HIMELSTEIN, P. (1966). Research with the Stanford-Binet, Form L-M: The first five years. *Psychol. Bull.*, 65, 156-164.

JONES, S. (1962). The Wechsler Intelligence Scale for Children applied to a sample of London primary school children. *Brit. J. Educ. Psychol.*, 32, 119-132.

MAXWELL, JAMES (1961). *The Level and Trend of National Intelligence.* London: Univ. London Press.

McDONALD, A. (1967). *Children of Very Low Birthweight.* London: Heinemann Medical Books.

RAVEN, J. C. (1958). *Guide to Using the Coloured Progressive Matrices, Sets A, Ab, B (Revised Order* 1956). London: H. K. Lewis.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION XXX (1949). *The Trend of Scottish Intelligence.* London: Univ. London Press.

TERMAN, L. M., and MERRILL, M. A. (1937). *Measuring Intelligence: A Guide to the Administration of the New Stanford-Binet Tests of Intelligence.* London: George G. Harrap.

TERMAN, L. M., and MERRILL, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M.* London: George G. Harrap.

THORNDIKE, E. L. (1921). Intelligence and its measurement. *J. Educ. Psychol.*, 12, 124-7.

VERNON, P. E. (1937). A study of the norms and the validity of certain mental tests at a child guidance clinic: Part I. *Brit. J. Educ. Psychol.*, 7, 72-88.

VERNON, P. E. (1938). *The Standardisation of a Graded Word Reading Test.* London: Univ. London Press.,

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw Hill.

(*Manuscript received* 22*nd August*, 1967)

THE TRISTAN DA CUNHAN CHILDREN'S CONCEPTS OF EQUIVALENCE

By A. G. DAVEY

(*Department of Sociology, University of York*)

SUMMARY. The learning and utilisation of categories is the most elementary and universal form of cognition by which man adjusts to his environment. In order to observe what discriminably different items the Tristans treated 'as if' equivalent and the strategies employed to render them equivalent a free grouping task was adopted which allowed the subjects to demonstrate their habitual mode of categorising. Groupings were classified by size, grouping strategy and criterial attributes. Two sets of observations were made, the first soon after the Tristans arrived in Britain and the second twelve months later. Developmentally, the first attempts at equivalence were made by abstracting the attributes which were sensorily most readily available. With increasing age there was a growing preference for grouping in functional and nominal terms. The transition appears to have been accomplished by the intermediary strategy of uniting a diversity of objects by means of an imaginative narrative. The place and function of thematising in cognitive growth are discussed.

I.—INTRODUCTION.

IN the late Autumn of 1961 the volcano on Tristan da Cunha erupted and the islanders were evacuated and brought to England via South Africa. How this unsophisticated in-bred community from a severely restricted environment would adapt to a complex technological society was a matter of considerable psychological interest. The Department of Psychology at University College, London, undertook to make a study of selected characteristics of the children of school age before adaptation took place and to make repeated assessments for a period of two years to observe any subsequent modifications. The investigation, under the direction of Gertrude Keir, was begun in March, 1962, and an account of the principal findings will be found in Keir (1965). The present study concerns one aspect of the children's cognitive development—the ability to form equivalence categories.

The task of reducing the complexity of the environment by rendering discriminably different objects and events equivalent is a universal one but since, as Bruner, Goodnow and Austin (1956) observe, the categories into which man sorts his world are inventions and not discoveries, they necessarily reflect the demands of the culture in which they arise, so that the learning and utilisation of categories is one of the principal means by which the child is socialised.

The essence of the classical approach to classification is that the subject should respond in some consistent manner to recurring features built into the experimental material and thus indicate that either he possesses or has learned the category which the experimenter has in mind. But any investigation into the cognitive processes of a cultural community about which little is known requires a method which will both allow the subjects to demonstrate their habitual modes of categorising and give the investigor some insight into how the categories are attained.

The work of Olver (Olver, 1961; Bruner and Olver, 1963; Olver and Rigney-Hornsby, 1966) and Rigney, (1962) provides such an approach. It has proceeded within the broad framework provided by Vygotsky (1962), but has departed from it significantly by analysing separately what they describe as

" the semantic and syntactic features of equivalence," i.e., the content and structure of groupings. The present study has kept fairly closely to Rigney's procedure, although some modifications were made both to the pictorial array used for the sorting task and to the classification scheme.

II.—PLAN OF THE INVESTIGATION.

(*a*) *The Subjects.*

When the investigation began there were forty Tristan children of school age. Three of these returned to the island with the advance party before the experimental procedure could be repeated twelve months later. Of the thirty-seven children for whom complete data are available, three boys and seven girls were attending infant school (mean age 6y. 4m.; mean IQ 93·9*), five boys and eleven girls were attending junior school (mean age 9y. 9m.; mean IQ 83·8), and seven boys and four girls were attending secondary school (mean age 13y. 1m.; mean IQ 79·4).

(*b*) *Material and Procedure.*

The subject is presented with an array of forty-two small coloured pictures of familiar objects. The pictures are always set out in the same order in six columns of seven. He is first asked to name each picture to ensure that he looks at every card. Then the experimenter says: " I want you to look at *all* the cards again. Take your time. Now what you have to do is to pick out the cards which belong together and put them on the table over there. There is no right or wrong way to do this. Any way you want to put the things together is fine as long as they belong together."

When the subject has made his choice the experimenter adds: " Spread them out on the table and as you put each one down tell me why it goes with the others."

In this way both the grouping strategy and the attributes utilised are revealed. Each subject is asked to make ten groups if he can. Every grouping and explanation is recorded in full so that it can subsequently be analysed by size, strategy and attributes.

(*c*) *Classification of grouping strategies and attributes.*

Three principal types of grouping strategy were distinguished: superordinates, complexive and thematic. Attributes have been divided into four categories: perceptible, functional, linguistic convention, and other. A full description of each type and sub-type of strategy and attribute is given in the Appendix, with an example extracted from the protocols.

III.—RESULTS.

The significance of the relationships between the various dependent variables and age and sex was obtained by means of a multiple regression analysis. (Regression coefficients are indicated in the text by 'b'). Figures 1 and 2, derived from this analysis, show the lines of best fit between dependent variables and age after the differential contributions of the sexes have been allowed for.

(*a*) *Size of Groupings.*

Any size of grouping between two and forty-two was theoretically possible but groups of more than six items were rare. Pairing was the preferred mode of grouping at each age level, despite the fact that the pictorial array offered a number of readily discriminable groups, e.g., there were six animals, five items of food and four articles of clothing.

---

* Revised Stanford-Binet Form L. For a detailed discussion of the children's performance on this and other intelligence tests, see Keir (1965).

While there was a reliable inverse relationship between pairing and increasing age ($b = -0{\cdot}46$ $p < {\cdot}05$) on the first testing this trend towards more inclusive groupings was not maintained on the re-test twelve months later. Nor was there any significant change in the means for pairing, although on the second test the girls did produce more groups of three or more than the boys ($b = -0{\cdot}3$ $p < {\cdot}02$).

In interpreting results from the Gelb-Goldstein-Weigl Scheerer Object Sorting Test, Goldstein and Scheerer (1941) consider that a decisive preference for pairing is indicative of extreme realism. The subject is unable to detach himself from the unique properties of objects. Yet, whereas approximately 50 per cent of the infant Tristans' pairings were based on some immediate perceptible attribute, only eleven (14 per cent) of the secondary children's pairs on the first occasion and four (5 per cent) on the second were so based. Furthermore, having made one pairing, say, on the criterion of ' brown-ness,' the young children were quite happy to go on to make a fresh pair on the same basis.

If the experimenter, like Goldstein and Scheerer, had been able to ask the maker of a pair if he could see anything to add to it or handed him other pictures and urge him to add them to the pair the basis of the strategy would have become more explicit, but at the same time would have biassed the subject's later groupings.

TABLE 1

PERCENTAGE OF DIFFERENT GROUPING STRATEGIES BY AGE AND SEX.

	First Year								
	Boys			Girls			All		
	I	J	S	I	J	S	I	J	S
Collection complexes . .	17	16	0	20	12	0	19	12	0
Other complexes	0	6	1	0	10	2	0	10	2
Themas	53	28	46	8	22	25	22	24	38
General superordinates	23	20	50	47	40	60	40	34	55
Itemized superordinates	7	22	3	22	12	13	17	15	5
Defeat	0	8	0	3	4	0	2	5	0
No. of strategies	30	50	70	70	110	40	100	160	110
	Second Year								
Collection complexes . .	24	18	1	3	1	0	9	6	1
Other complexes	0	6	3	0	5	0	0	6	2
Themas	3	30	56	7	22	10	6	24	39
General superordinates	63	32	30	77	61	78	73	53	47
Itemized superordinates	7	14	10	13	11	12	11	11	11
Defeat	3	0	0	0	0	0	1	0	0
No. of strategies	30	50	70	70	110	40	100	160	110

(*b*) *Grouping Strategies.*

Both Rigney's and Olver's results indicate that with age there is a decline in complexive groupings and a growing ability to keep the basis of equivalence invariant. The regression analysis failed to reveal any reliable relationships between grouping strategies and age in the present data, although there is a significant increase in the means for superordinates between 1962 and 1963 ($t=2{\cdot}42$, $p<{\cdot}05$). Table 1 suggests that on the second testing superordinates may actually decline with age, but the regression coefficient is not significant at the 5 per cent. level.

The most remarkable difference between the Tristans and the two American samples is in their use of thematic groupings. Olver's subjects produce one in a total of 840 responses. Rigney, using pictorial instead of verbal material, obtained a substantial number but the frequency decreased steadily with age. The Tristans, in contrast, appear to progress towards the thematising rather than away from it. The trend which is apparent in the first set of results (Table 1) is confirmed by those obtained twelve months later ($b=2{\cdot}47$, $p<{\cdot}02$).

This result is of interest since the thematising mode does not appear to have been favoured by Mexican children (Maccoby and Modiano, 1966) or by the Eskimo and Senegalese groups studied by Greenfield, *et al.* (1966).

TABLE 2

PERCENTAGE OF DIFFERENT ATTRIBUTES BY AGE AND SEX.

	First Year								
	Boys			Girls			All		
	I	J	S	I	J	S	I	J	S
Perceptible colour	10	0	0	58	27	18	43	18	7
Perceptible other	10	8	9	8	9	2	9	9	5
Functional egocentric	10	22	20	1	9	34	4	15	25
Functional objective	0	8	16	0	13	12	0	10	16
Linguistic	0	8	9	3	4	10	2	5	9
Other	17	16	0	19	10	0	18	12	0
No attribute	53	30	46	8	25	24	22	26	38
Defeat	0	8	0	3	3	0	2	5	0
No. of attributes	30	51	70	70	116	41	100	167	111
	Second Year								
Perceptible colour	33	2	0	45	14	0	41	10	0
Perceptible other	24	6	3	32	16	5	30	13	4
Functional egocentric	7	14	15	6	20	50	6	19	29
Functional objective	13	26	17	7	15	18	9	19	16
Linguistic	0	4	8	0	11	17	0	8	11
Other	17	18	1	3	1	0	7	6	1
No attribute	3	30	56	7	23	10	6	25	39
Defeat	3	0	0	0	0	0	1	0	0
No. of attributes	30	50	72	70	113	40	100	163	112

N.B.—The number of attributes for some columns is greater than that for the respective column in Table 1 since some complexes required a multiple classification.

(c) *Attributes.*

The attributes used by the three age groups as a basis for their equivalence groupings are set out in Table 2. For the purposes of the multiple regression anaylsis all the classifications employing surface qualities, colour, shape, etc., have been pooled and labelled perceptible, and the functional and linguistic categories have been pooled and labelled non-perceptible. The responses in the remaining categories, rhyme, affect, etc., were too sparse to merit any attempt to establish age trends or sex differences.

The results for both years (Table 2) indicate that growth brings a progressive independence from the surface qualities of things as a basis for equivalence ($b = -0{\cdot}45$, $p < {\cdot}05$ and $b = -0{\cdot}93$, $p < {\cdot}001$) and an increasing reliance on non-perceptible attributes ($b = 0{\cdot}61$, $p < {\cdot}01$ and $b = 0{\cdot}72$, $p < {\cdot}001$). This is substantially in accord with the American data, but there are three notable differences. Whereas Rigney's subjects used form almost as often as they used colour the Tristan's perceptible groupings were made predominantly on the basis of colour. When he does categorise by shape the Tristan tends to respond to the total gestalt. Only occasionally were the internal details of the pictures treated as criterial attributes. Secondly, the Tristans appear to be slow to grasp the possibilities of functional attributes and ready-made lexical categories for transforming the kaleidoscope of sensory impressions into equivalence units. Rigney's 6-year-olds employed non-perceptible attributes for 36 per cent. of their groupings compared with the Tristan infants' 6 per cent. on the first occasion and 15 per cent. on the second. Thirdly, while linugistic classifications are not highly represented at any level amongst the Tristans, they account for almost a fifth of the Bostonians' groupings at age 8 and a third of their groupings at 11.

IV.—DISCUSSION.

The Tristan data exhibit the same perceptible-functional trend as that of the more sophisticated Bostonians. But whereas the American child's first attempts at " packaging properties into non-perceptible units of belonging " (Bruner and Olver) appear to be in terms of what he can do to things, Figures 1 and 2 suggest that the Tristan requires the intermediary step of thematising before he can fully exploit the combinational possibilities of non-perceptible attributes.

Although thematic type groupings have generally been regarded as elementary or regressive strategies (e.g., Kagan, Moss and Sigel, 1963 ; Goldstein and Scheerer, 1941) this does not necessarily mean that the Tristans are arrested at an early point of the concrete-abstract continuum since it is possible that thematising is the product of developing discrimination, and the precursor of a more efficient application of the rules of superordination.

Superordination is evident well before the Tristan child is able to disengage himself from the immediacy of his sensory impressions. His generalisations are made possible by the coarseness of his differentiation. He seizes on some sensorily prominent feature and remains comparatively unaware of a whole range of within category differences. As his ability for finer discrimination develops he becomes increasingly aware of the unique concreteness of things and of the inadequacies of perceptible criteria for reducing the new multiplicity of identities to manageable units. But this growing capacity to explore the specificity of things is not necessarily matched by either the ability or the vocabulary for generating subtler strategies for achieving equivalence. One way out of this dilemma is to unite things in the manner of the constant conjunctions he experiences in his daily life : boats, boots and fish, or penny, bike and sweets. It is not an economical encoding since a thema is not less than the sum of its parts, but it does serve to release the perceiver from the vivid individuality of things until he can extend the superordinating strategy beyond mere perceptual inclusion.

The present data are, of course, too sparse and too unrepresentative to substantiate this argument for the transitional function of thematising, although some support can be derived from Annett (1959), who found that grouping explanations in terms of physical contiguity were curvilinear with age. If,

Fig. 1 **FIRST YEAR**

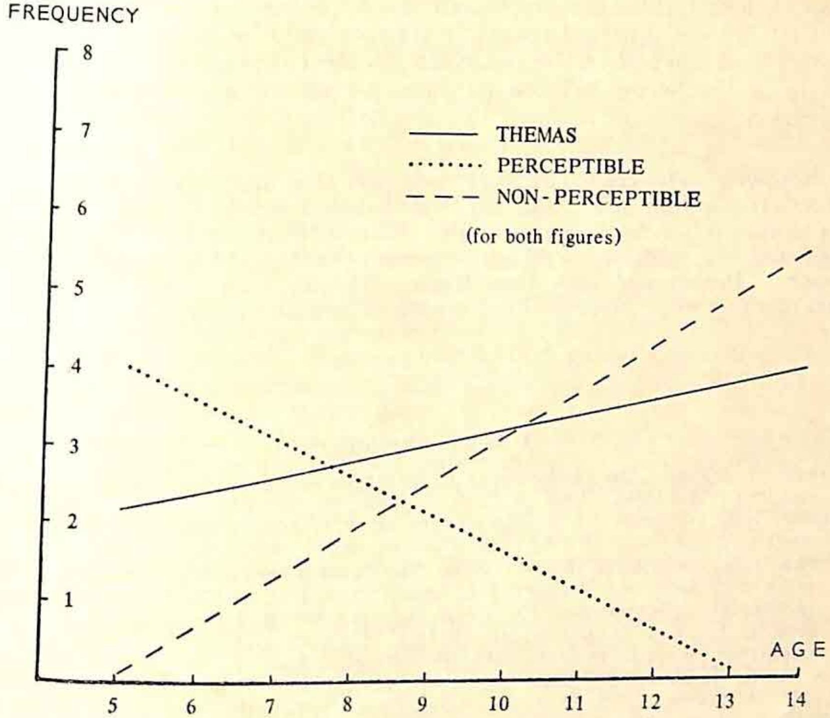
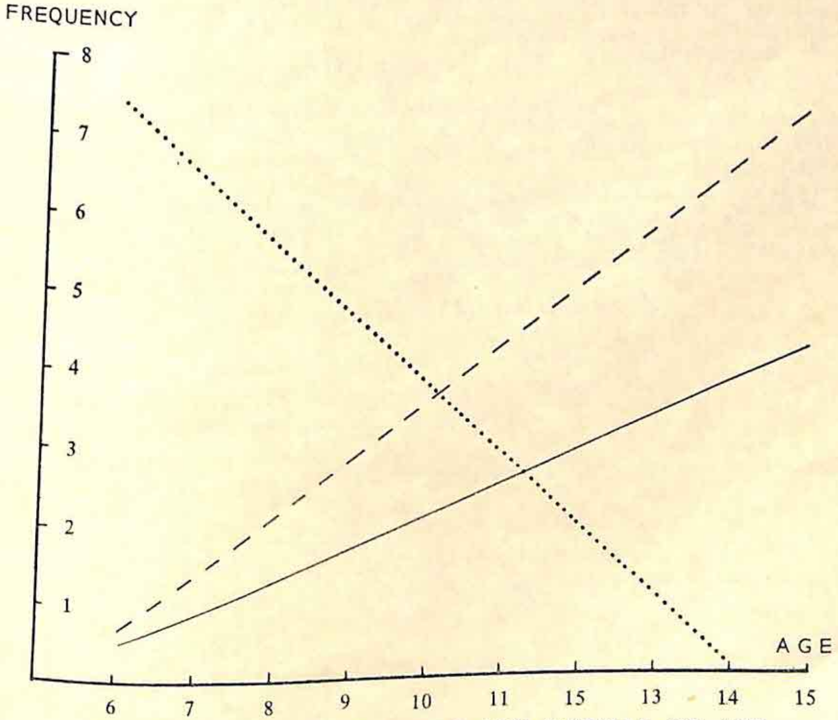


Fig. 2 **SECOND YEAR**



REGRESSION OF ATTRIBUTES AND THEMAS ON AGE however, a longitudinal study showed the hypothesis to be valid, it would imply that quite a different cognitive status should be accorded to thematic groupings from the rest of the syndrome of object dependent responses which Goldstein and Scheerer take to be characteristic of a concrete attitute to classification.

ACKNOWLEDGEMENTS.—This study was part of a larger investigation directed by Miss Gertrude Keir and financed by the Medical Research Council. The author wishes to record his indebtedness to Miss Keir, without whose authority and encouragement this research would not have been possible, and his grateful thanks to Professor J. Bruner and Miss Joan Rigney (Hornsby) for making available the original pictures which were modified for the classification test.

V.—REFERENCES.

ANNETT, M. (1959). The classification of instances of four common class concepts by children and adults. *Brit. J. Educ. Psychol.*, 29, 223-236.

BRUNER, J. S., GOODNOW, J. J., and AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

BRUNER, J. S., and OLVER, R. R. (1963). The development of equivalence transformations in children. In WRIGHT, J. C., and KAGAN, J. (Eds.), Basic cognitive processes in children. *Monogr. Soc. Res. Child Development*, 28, 2, Serial No. 86.

GOLDSTEIN, K., and SCHEERER, M. (1941). Abstract and concrete behaviour: an experimental study with special tests. *Psychol. Monogr.*, 53, 2.

GREENFIELD, P. M., REICH, L. C., and OLVER, R. R. (1966). On culture and equivalence: II. In BRUNER, J. S., *et al. Studies in Cognitive Growth*. New York: Wiley.

KAGAN, J., MOSS, H. A. and SIGEL, I. (1963). Psychological significance of styles of conceptualization. In WRIGHT, J. and KAGAN, J. (Eds.), *op. cit.*

KEIR, G. (1965). The psychological assessment of the children from the island of Tristan da Cunha. In BANKS, C. and BROADHURST, P. L. (Eds.). *Studies in Psychology*. London: Univ. of London Press.

MACCOBY, M., and MODIANO, N. (1966). On culture and equivalence: I. In BRUNER, J., *et al., op. cit.*

OLIVER, R. R. (1961). *A Developmental Study of Cognitive Equivalence*. Unpublished doctoral dissertation. Cambridge, Mass.: Radcliffe College.

OLVER, R. R., and RIGNEY-HORNSBY, J. (1966). On equivalence. In BRUNER, J., *et al., op. cit.*

RIGNEY, J. C. (1962). *A Developmental Study of Cognitive Equivalence Transformations and their use in the Acquisition and Processing of Information*. Unpublished honours thesis. Cambridge, Mass.: Radcliffe College.

VYGOTSKY, L. S. (1962). *Thought and Language*. New York: Wiley.

(*Manuscript received* 13*th September*, 1967)

VI.—APPENDIX.

CLASSIFICATION OF GROUPING STRATEGIES

(1) SUPERORDINATE GROUPINGS.

(i) *General superordinates.* Items are grouped on the basis of one or more attributes common to them all, e.g., saw, scissors, hammer. "They are all tools."

(ii) *Itemised superordinates.* Items are linked by a common property and this is explicitly stated for each item in turn, e.g., apple, orange. "The apple grows on a tree and the orange grows on a tree."

(2) COMPLEXIVE GROUPINGS.

A group is formed by selecting a variety of attributes none of which is common to all the items. This strategy may take one of five forms:

(i) *Collections.* Different attributes are specified for each item but the subject fails to abstract any superordinating quality, e.g., hammer, nails. "With a hammer you hit nails down and a saw saws wood."

(ii) *Chain complexes.* There is no common attribute but each item is related to the one which precedes it, e.g., apple, orange, carrots. "An apple and an orange are both round and the orange is the same colour as the carrots."

(iii) *Key ring complexes.* A central item is related by some attribute to each of the other items but there are no explicit links between the subordinate items, e.g., pie, orange, apple, knife, fish. "The knife can cut an orange. The knife can cut an apple, a pie and a fish."

(iv) *Association complex.* The subject establishes an association between two items and then uses this association as the bond for additional items, e.g., tree, saw, garage, barn. "The saw can cut down the tree. The wood can make a garage and the wood can make a barn."

(v) *Multiple grouping.* Two or more sub-groupings are formed on the basis of different attributes and often with different structures but the subject fails to abstract any superordinating attribute which will bridge the gap, e.g., hammer, nails, shoes, bicycle, saw, screw, ruler. "The hammer, nails and shoes belong together because the little nails go in the shoes. When a bike is broken, it is dangerous. You can measure it with a ruler and saw something off and screw something in. You can cut something off it with a knife."

(3) THEMATIC GROUPINGS.

The items are not united by a common bond, but are held together by the child's imaginative account, which may be a single sentence or a short narrative, e.g., clock, cow. "The clock tells you when it is time to milk the cow." Or, bicycle, doll, penny, Wellington boots. "The girl is riding her bike and she has a penny to spend. She has to put on her seaboots because it is a wet day."

CLASSIFICATION OF ATTRIBUTES

The attributes which the subjects selected as the basis of equivalence have been divided into four broad groups with sub-divisions.

(1) PERCEPTIBLE ATTRIBUTES.

The items are matched by virtue of some surface quality or are spatially related to one another.

(i) *Colour.* e.g., flower, lamp. "That's yellow and that's yellow."

(ii) *Shape.* e.g., clock, penny. "They're both round."

(iii) *Analytical.* The subject ignores the general outline of the object and selects some detail as the basis of equivalence, e.g., tap, scissors. "They've both got a little screw in them."

(iv) *Spatial.* e.g., flower, lamp, tap, candle, house. "They all belong in the house."

(2) FUNCTIONAL ATTRIBUTES.

(i) *Functional egocentric.* Items are equated in terms of what can be done to the objects or in terms of what they can do for the subject, e.g., apple, orange. " I can eat an apple and I can eat an orange."

(ii) *Functional objective.* Here the function is intrinsic to the objects, e.g., flowers, tree. " They both grow."

(3) LINGUISTIC CONVENTION (nominal).

The subject explains his grouping by giving a class name. " They are all animals " (tools, fruit, etc.).

(4) OTHER TYPES OF ATTRIBUTES.

(i) *Affective.* The items are rendered equivalent either in terms of the emotion they arouse—" I like them both "—or in terms of some internal state which the subject ascribes to the items themselves, e.g., apples, nails, candle, boat. " These two go together because they are happy together, but they don't like these two." This type of response was, in fact, confined to one very disturbed boy.

(ii) *Rhymes.* e.g., tree, bee. " They rhyme."

(iii) *Fiat.* The subjects assert that the items are the same or different but is unable to give the basis of the relationship, e.g., apple, orange. " That's an apple and that's an orange. Well, they're nearly the same really."

(iv) *Defeat.* The subject declares that he is unable to find any more cards which go together.

(v) *No attribute.* This entry on Table 2 refers to the thematic groupings. When the subject adopts this strategy, items are incorporated into a group because he considers them appropriate to the situation he has created, but the salient attributes cannot be determined with certainty.

TEACHING APPROACH AND THE DEVELOPMENT OF DIVERGENT THINKING ABILITIES IN PRIMARY SCHOOLS

By F. A. HADDON
(*Rolle College, Exmouth*)

AND HUGH LYTTON
(*University of Exeter Institute of Education*)

SUMMARY. This research was designed to evaluate the effects of differing teaching approaches on divergent thinking abilities, the hypothesis being that informal, progressive teaching would promote these abilities more than formal, more subject-centred teaching. Two-hundred-and-eleven children, 11 to 12 years old, of all ability levels, were tested, half coming from 'Formal' and half from 'Informal' primary schools which were matched for verbal reasoning quotient (VRQ) and socioeconomic background. The results showed that pupils from the Informal schools were significantly superior in divergent thinking abilities. The difference in means often reached the ·01 level of significance and was in no case in the unpredicted direction. The hypothesis was thus confirmed.

A subsidiary hypothesis—that the correlation between VRQ and divergent thinking abilities would decrease as the VRQ of each sub-group rose—was also supported. As expected, these correlation values were generally higher in the Informal schools.

I.—INTRODUCTION.

DESPITE the attention which 'Creativity' has received as a fashionable topic for research there is, as yet, as Vernon (1964) and Hudson (1966) point out, no clear evidence for the assumption that high scorers on tests of 'divergent thinking' in Guildford's terminology, will be particularly fertile in creative original production in their own life situation. Nor is there evidence for the converse, that high scorers on tests of 'convergent thinking' alone, will lack such creativity. The best way of looking at this question probably is to regard high scores on divergent thinking tests as an indication of lack of anxiety about non-conformist responses, a necessary, but not a sufficient condition for creative work. To avoid undue claims, we have used the term 'divergent thinking,' rather than 'creativity' throughout.

As Guilford has shown in his many studies (summarised in Guilford 1956) divergent abilities can be distinguished factorially from convergent abilities which have hitherto formed the main component of general ability tests. It is important to explore divergent abilities since they hold some promise of providing new insights into cognitive functioning. Tests of these abilities may provide a welcome supplement to 'intelligence tests' as we have so far known them, and indicate the type rather than the level of future performance (Hudson, 1966). This model of cognitive abilities implies that divergent and convergent thinking are complementary aspects or different styles of intellectual functioning.

Whilst a great deal has been written on the guidance of creative ability, very little work appears to have been published investigating the effects of particular school situations. Torrance does hint briefly at the effects of school orientation towards learning (Torrance, 1962) and Hasan and Butcher (1966) report his view to the effect that the success of predominantly divergent thinkers is directly related to the degree of freedom and permissiveness and the lack of authoritarian discipline within a school. Vernon (1964), in a review of creativity and intelligence, also comments : " I strongly suspect some schools do much more to stimulate and foster, or else to inhibit, creative talent than others." Hilgard and Sears (1964) quote Spalding (1963) as finding strong negative relations between the expression of creativity in elementary-aged children and teacher behaviour characterised as formal group instruction, using shame as a punishment technique, whilst Sears (1963) found positive correlations between creativity and teachers' use of the technique of rewarding children by personal interest in their ideas rather than by evaluation.

The research reported here was concerned with evaluating the effects of contrasted Formal and Informal schools upon performance on tests of divergent thinking abilities. It was felt that teaching approach might be a variable affecting the degree of divergency exhibited in children's thinking. Two other aspects have also been examined. First, the relationship between IQ and performance on divergent tests has been looked at in some detail. Secondly, the relationship between performance on divergent tests and peer popularity has been compared with the relationship between VRQ and peer popularity in each of the schools used.

II.—BASIC HYPOTHESIS AND RESEARCH DESIGN.

It is generally accepted that abilities and personality characteristics are, to a large extent, moulded by environmental influences such as family and school. Since the school, however, is only one factor in the development of a child's personality, the basic structure of which has very largely been laid down by the time the child arrives there, measurable differences between school and school are likely to be small. The basic hypothesis is that some differences in divergent thinking abilities can, nevertheless, be detected between comparable children who have spent formative years in contrasted environments such as different primary schools provide. Schools may differ in many ways, one of which is the degree of formality or informality which permeates the approach to learning. The two types of schools which have been contrasted here are the Formal, or traditional, school which places emphasis upon convergent thinking and authoritative learning and the Informal or progressive school, where the emphasis is upon self-initiated learning and creative activities. It was predicted that children from the Formal school will have lower mean scores on ' divergent ' tests than children from the Informal school, which, it was thought, encourages the growth of personality traits associated with divergency.

A further area of interest was the relation between IQ or VRQ (verbal reasoning quotient) and measures of divergent thinking. It was thought that given a high general cognitive potential, development along either convergent or divergent lines may receive considerable emphasis at the expense of the other, but that the Informal school, more than the Formal school, would develop both convergent and divergent thinking in line one with the other. Conversely, when intelligence is limited, VRQ and divergent thinking ability are likely to display greater correlation. In other words, the differentiation of divergent from convergent abilities is likely to show itself more clearly at higher than at lower levels of general intellectual ability (whether measured via convergent or divergent abilities). This prediction is in accordance with Yamamoto's (1965) findings. To summarize the predictions : (1) Mean scores on divergent tests would be significantly higher in the Informal schools than in the Formal schools. (2) Correlation between VRQ and divergent thinking abilities would decrease as the mean VRQ and mean divergent tests scores of sub-groups rose, but the values obtained would throughout be higher in the Informal school.

The Tests.

Full details of the tests and the system of scoring, adapted from that used by Torrance for the Minnesota Tests of Creative Thinking, are given in the Appendix. Briefly, the tests were :

The Non-verbal or 'Iconographic' Tests.

(1) Circles (Torrance, 1962).

(2) Vague shape of dots (developed from Torrance, 1962, Picture Construction Task).

(3) Block-printing (New test).

The Verbal Tests.

(4) Uses for a shoe-box (modified from Uses for a Can, Torrance, 1962).

(5) Problems which might arise in taking a bath (Torrance, 1962, Common Problems).

(6) Imaginative stories (Torrance, 1962).

(7) The Camping expedition, a disguised sociometric test (New test).

From each test two scores were obtained, one of which was Originality (except for Test 2—see Appendix). The total for each test was the total of the two component scores.

The Subjects and Schools.

The schools used were selected after consultation with lecturers from a college of education who were familiar with the area, and after advice from a local inspector of schools. Two pairs of contrasted primary schools matched for socio-economic background were chosen. A and C are the Formal, B and D the Informal schools. Schools A and B are in a predominantly middle-class urban area, the mean VRQ being 106·5 and 103·5, respectively. Schools C and D draw their children from a more mixed social background in a different urban area, the mean VRQ being 96·1 and 98·3, respectively. The mean VRQ of the Formal schools combined (A and C) is 101·75, and that of the Informal schools (B and D) 101·14.

The tests were administered to all the children eligible for transfer to secondary schools the following term. In all, 211 children, 11–12 years old, were tested and the whole ability range in the schools was covered.

There were several advantages in choosing to administer the tests to children with;n a few days of completing their primary education :

(i) Such children had been subjected to the effect of their particular school for the maximum amount of time.

(ii) Selection for secondary education had already taken place, so that the children were not concerned about possible effects upon their future schooling.

(iii) IQs (or VRQs) for all the children were made available by the Chief Education Officer. These were based on Moray House Verbal Reasoning Test No. 72, which was used throughout the County as part of the selection procedure for allocation to secondary education.

III.—RESULTS AND DISCUSSION.

Comparisons between two types of school.

The results of the tests are set out in Table 1. Each component was marked separately and the component scores combined to obtain a test score for each child. Mean scores for the group were calculated from these.

TABLE 1

COMPARISON OF MEAN TEST SCORES (RAW SCORES).

	FORMAL SCHOOLS			INFORMAL SCHOOLS			t-value of difference Combined Means	P
	A Mean	C Mean	Combined Mean A+C	B Mean	D Mean	Combined Mean B+D		
VRQ ..	106·50	96·1	101·75	103·50	98·30	101·14	0·31	N.S.
Test 1	12·00	9·1	10·55	15·55	10·10	12·80	3·07	< ·01
2	5·72	5·92	5·82	7·48	7·52	7·50	5·48	< ·01
3	7·50	4·88	6·19	9·05	11·64	10·35	6·50	< ·01
4	4·98	6·58	5·78	6·86	7·62	7·25	2·11	< ·05
5	10·56	10·08	10·32	15·20	8·41	11·81	1·61	N.S.
6	12·46	12·50	12·48	14·75	13·54	14·15	2·51	< ·05
N	54	50	104	59	48	107		
N for Test 3—(numbers reduced on account of practical testing difficulties).								
	29	25	54	35	25	60		

The results in Table 1 in general confirm the main prediction of significantly higher scores being obtained from the Informal schools. Even in Test 5, where the difference does not reach statistical significance at the ·05 level, the direction of the difference is in line with the other results. There is, of course, no appreciable difference in mean VRQ between the two groups of schools.

The Iconographic Tests—numbers 1, 2 and 3—all show highly significant differences between the two groups. This is not unexpected. It is reasonable to suppose that the effect of the different approaches will show to its maximum here, in that no child will be handicapped by having to respond in writing. This not only allows the less able child to make a response but applies particularly to the boys who often seem either unable or reluctant to express themselves in writing with the same fluency as girls of similar age. Conversely, one would expect lesser differences between the two groups on the Verbal Tests. This is generally true again, Tests 4 and 6 showing a difference of means significant at the ·05 level, whilst Test 5 just fails to reach this level.

TABLE 2

COMPARISON OF MEAN SCORES OF TEST COMPONENTS (STANDARD SCORES).

Component	Tests from which obtained	Formal Schools Means	Informal Schools Means	't' value of difference	P
Fluency	6	7·26	7·67	1·78	N.S.
Flexibility	1, 4, 5	57·31	61·39	2·62	< ·01
Originality	1, 4, 5, 6	61·61	70·10	2·71	< ·01
Fit of Concept	2	3·26	3·81	3·12	< ·01
Elaboration	2	2·58	3·69	5·30	< ·01

Having confirmed the main prediction, further analyses were made. The component raw scores which were used for the previous table were changed to standard scores with a mean of 20 and a standard deviation of 5 to make them comparable. It was then possible to extract the component scores from each test to obtain a total component score for Originality, etc., for the whole battery. The results are shown in Table 2. Scores for Test 3 were not included here because only 50 per cent of the children were tested.

An alternative way of looking at the results is shown in Table 3, in which the scores were added to give mean scores for Iconographic Tests, Verbal Tests, and all Divergent Tests. Comparisons are made over different ranges of VRQ. Results from Test 2 were not included in these calculations since the test was scored in a somewhat experimental manner.

TABLE 3

COMPARISON OF MEAN SCORES OVER VARIOUS RANGES OF VRQ (STANDARD SCORES).

	Formal Schools Means	Informal Schools Means	't' value of difference	P
Over full range of VRQ:				
VRQ	101·75	101·14	0·31	N.S.
Iconographic	18·96	21·29	4·16	< ·01
Verbal	19·50	20·72	2·71	< ·01
All Tests	38·34	41·87	4·15	< ·01
VRQ 100 and above				
VRQ	113·47	110·57	1·88	N.S.
Iconographic	20·18	22·37	2·88	< ·01
Verbal	20·29	21·97	3·18	< ·01
All Tests	40·36	44·18	3·62	< ·01
VRQ below 100:				
VRQ	88·59	89·11	0·32	N.S.
Iconographic	17·59	19·92	3·11	< ·01
Verbal	18·61	19·13	0·76	N.S.
All Tests	36·06	38·92	2·41	< ·05

The results again gave a clear indication of the superiority of the Informal schools in performance on the tests. It seemed important next to examine the differences between the two types of school over high and low ranges of VRQ. Whilst it is recognised that the VRQ is in some ways an unsatisfactory indication of general intelligence because it samples too narrow a range of intellectual abilities, nevertheless, considerable importance is attached to this index and to the abilities it measures, and it usually carries considerable weight in selection procedures. If, as has been claimed, the Informal schools are fostering the development of divergent thinking abilities, then one would expect a greater difference between the two types of school over the higher ranges of VRQ where there is greater intellectual potential. Conversely, over the lower ranges of VRQ, the expected differences would be less. Table 3 compares the two types of school over the range VRQ 100 and above, and VRQ below 100, respectively.

The results shown in the lower part of Table 3 support the main hypothesis, but the prediction of greater differences with higher VRQs is borne out only as regards the verbal, not the iconographic tests. It is worth noting that for VRQs above 100 the differences are in the reverse direction to the differences on the divergent tests, yet in spite of this, the Informal schools achieve results which are significantly superior.

The Evidence from Correlations.

The main prediction with regard to correlation, it will be recalled, was that the correlation between VRQ and divergent thinking abilities would decrease as VRQ rose. This is demonstrated by Table 4 which used the combined data from the four schools. The results here are in agreement with Yamamoto (1965) who reports " a consistent decrease in the size of correlation as the IQ level of sub-groups became higher " (although the values quoted by him are considerably smaller in both the school systems used for his investigations than the values obtained here.) In the lower half of Table 4 the data have been re-grouped on the basis of total scores on the tests of divergent thinking.

TABLE 4

CORRELATION OF VRQ AND DIVERGENT THINKING TESTS.

ALL SCHOOLS. Sub-groups based on VRQ ranges.	N	r
Full range of VRQ 70–135	211	·480
VRQ 115 and above	35	·076
VRQ 100 and above*	115	·164
VRQ below 100	96	·512

* This includes the category VRQ 115 and above.

ALL SCHOOLS. Sub-groups based on Divergent Thinking Scores. Mean Divergent Thinking Score=40	N	r
Full range	211	·480
Divergent Score 40 and above	103	·230
Divergent Score below 40	108	·440

These findings strongly confirm the view that convergent thinking and divergent thinking are two complementary aspects of intellectual ability in general (styles of functioning). They separate from each other, each attaining a more clearly independent status of its own, at higher levels of ability, and they overlap more closely at lower levels, whether ability is measured by convergent or by divergent tests.

Within this general concept of a diminishing relationship at higher levels of ability, a subsidiary hypothesis was that the correlation would be somewhat higher in the Informal schools. It was argued that in these schools, because of their particular approach to learning, the link between the two aspects would be maintained to a greater extent. The idea is more forcibly expressed by saying that formal education will tend to destroy this connection by putting a premium on convergent thinking and conformist behaviour. Table 5 explores this connection over various VRQ groupings.

TABLE 5

CORRELATIONS BETWEEN VRQ AND DIVERGENT THINKING TESTS.

		Formal Schools				Informal Schools		
VRQ	N	Iconographic r	Verbal r	All Tests r	N.	Iconographic r	Verbal r	All Tests r
Full Range	104	·335	·418	·454	107	·379	·618	·560
115 and above	22	—·041	—·011	—·050	13	·018	·026	·036
100 and above *	55	—·051	·159	·059	60	·204	·483	·366
Below 100	49	·245	·576	·487	47	·367	·578	·548

*This includes the category 115 and above.

The evidence in this table is quite convincing in its support of the hypothesis In every instance the value of *r* is higher in the Informal schools. This is especially noticeable in the range VRQ 100 and above, which suggests that formal education fails to develop latent ability to think along divergent lines among the most able children.

It is interesting to compare these results with the outcome of a pilot study in Birmingham (Kellmer Pringle and McKenzie, 1965) which involved two contrasted primary schools selected on criteria similar to those used in this study. The object of the research was to examine teaching methods as a factor influencing the degree of rigidity in problem solving. The definition of rigidity as " the inability to restructure a field in which there are alternative solutions to a problem in order to solve that problem more efficiently," is almost a negative definition of divergency and it might have been expected that the Informal school would have shown significantly less rigidity among its pupils. In fact, no overall difference was found, but since the comparison was made on the degree of set found in performance on one arithmetic test, this is hardly sufficient evidence for any definite conclusions to be drawn.

Sociometric Analysis.

Considerable attention has been given in America to the social relationships of creative individuals—in our terminology, those who do well on tests of divergent thinking abilities. The general view seems to be that such children are subjected to pressure to conform and to conceal their divergent characteristics. It is suggested that the creative child will be lacking in popularity and that social pressures may well be a factor inhibiting the development of divergent abilities (Torrance, 1962).

However, it was felt that the qualities associated with the divergent personality—fluency, flexibility and sensitivity—might well be qualities which make for good social relationships in a society in which originality was valued, and that there could, therefore, be a positive relationship between popularity and divergent personality. Test 7 was included in the battery to explore this connection. It was expected that any correlation found would be low, but that the values would be higher in the Informal schools. The argument here was that in so far as a school is able to transmit its cultural values, those values should be reflected in the general appreciation of the individuals who exhibit them in action. As a corollary to this, the correlation between VRQ and popularity should show a higher value in the Formal schools.

Table 6 sets out these correlations. These have been calculated on a school by school basis because the value of sociometric status as expressed here varies according to the number of children in the class, and so it was not possible to combine the schools for comparisons. Spearman's rank order correlation has been used.

TABLE 6

Rank Order Correlation Coefficients Between VRQ and Sociometric Status, and Between Divergent Thinking Scores and Sociometric Status.

School	N	VRQ Range	VRQ and Sociometric Status	Divergent Thinking and Sociometric Status
A. Formal	33	135—100	·137	—·128
C. Formal	50	126—70	·337	·279
B. Informal	47	135—85	·058	·363
D. Informal	37	131—71	·462	·382

There is partial support for the prediction, but two values require some discussion. In School A the top class contained only children of VRQ 100 and above. The probable result of this is to give a much lower value for both correlations. However, the important point here is not so much the value of the coefficients as the relationships between them. VRQ has a positive correlation with sociometric status, whereas divergent test scores have a negative one—as they did in Getzels' and Jackson's study.

The second point to discuss is the values obtained from School D. The three other schools all show the predicted direction in the differences between the two values. This is not only reversed in School D but the values themselves are the highest obtained. This school is a particularly delightful one to visit and gives one the impression of friendly informality combined with high standards of work and behaviour. It may be that the ethos of this school ensures esteem for both types of ability and thus makes for popularity of both high convergent and high divergent thinkers, who indeed are often the same children.

The overall results suggest that the school ethos as formulated by the head and the staff plays an important role in determining the children's values during the primary stage. This would seem to explain the inter-school differences in our investigations.

IV.—Conclusions.

The analysis of the test results provided considerable evidence in support of the chief hypothesis. This was, to repeat it briefly, that the Informal schools provide an environment which develops qualities of personality that result in a high level of divergent thinking ability. This has been supported by comparisons of mean test scores classified in a variety of ways. The means from the Informal schools have in most cases been significantly higher than those from the Formal schools. In those cases where the difference in means has not reached the level of statistical significance the direction of the difference has confirmed the general trend. In *no* case has there been a difference in the unpredicted direction.

The evidence from correlations was again in support of the contention that the Informal school succeeded in developing more of the divergent potential, since the correlation coefficients between VRQ and divergent test scores were consistently higher in the Informal schools. The general hypothesis of a decrease in the correlation value as both the VRQ and the divergent ability level of sub-groups rose was also confirmed.

The results from the sociometric analysis are inconclusive, but there is some indication that there is a positive connection between divergent ability and peer popularity in the Informal schools. This differentiation appears to increase if the lower VRQs are excluded (Schools A and B).

This study has not been comparing good *versus* bad schools, but good schools which operate with a somewhat different emphasis. It is certainly not permissiveness which is the distinguishing criterion of Informality as opposed to Formality. The most striking difference lies in the degree of emphasis laid upon self-initiated learning. Behind this emphasis in the Informal schools and fundamental to its success, lies the pattern of inter-personal relationships within the school. One's impression in the Informal school is of a relaxed, friendly atmosphere in which children move freely, both within the classroom and in the school generally. Particularly noticeable is the freedom of access to the libraries and the extent to which children work in them unsupervised. The Formal schools are not unfriendly but one senses a tighter rein and a firmer directive. Class work is more in evidence.

The investigation has shown that the optimum development of divergent thinking abilities is related to a certain teaching approach. As to what lies at the roots of this relationship we can only speculate. But it would seem that it is based on the teacher's confidence in the child's ability to think adventurously and in new directions, which in turn, will determine the child's estimation of himself and of his abilities. If the teacher can enter into the child's thinking, if she is prepared to let work develop in unexpected directions according to the child's needs and interests, if she can find and express genuine pleasure in the child's efforts, then self-initiated learning can be developed. It is in this climate that divergent thinking abilities are seen to flourish.

V.—REFERENCES.

ANDERSON, J. E. (1960). The nature of abilities. In E. P. Torrance (Ed.), *Education and Talent.* Minneapolis, Minnesota : Univ. of Minnesota Press, pp. 9-31.

GUILFORD, J. P. (1956). The structure of intellect. *Psychol. Bull.*, 53, 267-293.

HASAN, P., and BUTCHER, H. J. (1966). Creativity and intelligence : a partial replication with Scottish children of Getzels' and Jackson's study. *Brit. J. Psychol.*, 57, 129-135.

HUDSON, L. (1966). *Contrary Imaginations.* London : Methuen.

KELLMER PRINGLE, M. L., and MCKENZIE, I. R. (1965). Teaching method and rigidity in problem solving. *Brit. J. Educ. Psychol.*, 35, 50-59.

SPEARS, P. S., and HILGARD, E. R. (1964). The teacher's role in the motivation of the learner. In *Theories of Learning and Instruction*, 63rd N.S.S.E. Yearbook. Chicago : Univ. of Chicago Press.

SPAULDING, R. (1963). *Achievement, Creativity and Self-Concept Correlates of Teacher-Pupil Transactions in Elementary Schools.* Urbana, Illinois : Univ. of Illinois (U.S. Office of Education Co-operative Research Project No. 1352) (mimeographed).

TORRANCE, E. P. (1962). *Guiding Creative Talent.* Englewood Cliffs, N.J. : Prentice-Hall.

VERNON, P. E. (1964). Creativity and intelligence. *Educ. Res.*, 6, 163-169.

YAMAMOTO, K. (1965). Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. *Brit. J. Educ. Psychol.*, 35, 300-305.

VI.—APPENDIX.

THE TESTS OF DIVERGENT THINKING ABILITIES AND SCORING PROCEDURES.

Unless otherwise indicated, the tests are all taken from the Minnesota Tests of Creative Thinking (Torrance, 1962). The scoring procedures are adapted from the same source. The titles of the Minnesota tests are given in brackets.

Test 1—*The Circles Test* (Circles and Squares Task).

Scoring—Flexibility : one point for each category of response.
Originality : one point for each response outside listed common responses.

Test 2—*The Vague Shape of Dots Test* (Derived from Picture Construction Task).

Scoring—Elaboration : one point for each new idea added to the initial response.

Fit of Concept : this was the new category of score which attempted to score for sensitivity to the qualities of the stimulus on a scale of 1–5 points.

Test 3—*The Block-printing Test* (New Test).

In this test the children were given an ink pad and a small block of wood, 2-in. × 1¾-in. × ½-in. and a booklet of absorbent paper. They were required to discover how many different kinds of mark they could make and to use these as they liked. (There were twenty-six possible types of mark.)

Scoring—Flexibility : one point for each category of mark discovered, e.g., long edge print, short-side print, etc.

Originality : one point for each image, each pattern, each textural experiment or for making letters or figures.

Test 4—*Uses for a shoebox* (Unusual Uses).

Scoring Flexibility— : one point for each category of response.

Originality : on the following basis : twelve or more children giving responses, score 0 ; six to eleven children giving responses, score 1 ; three to five children, score 2 ; two children, score 3 ; one child, score 4.

Test 5—*Problems that might arise in taking a bath* (Common Problems).

Scoring—Flexibility : one point for each response. (In this test each response constituted a new category).

Originality : As for Test 4.

Test 6—*Imaginative Stories* (Imaginative Stories).

Scoring—Fluency : points awarded for length on a scale of 1– 3.

Originality : one point for each of the following characteristics if they appeared at all : Description, vividness, indication of feelings, personal involvement, original solution or ending, original plot or setting, humour, inventive language, other original twists not covered by previous categories.

The scripts were scored by three judges, two of whom were unaware which schools were classed as Formal and Informal. The scores were then averaged for Fluency and Originality.

Test 7—*The Camping Party Test* (New Test).

This was a simple sociometric test disguised as another test of divergent thinking ability. A simple sociometric status was obtained by counting the number of nominations to join the expedition which each child received. This information was not obtained for all the children who were tested since in three of the schools used, the children of lower ability were spread through other classes from which they were extracted for testing purposes. This meant that they were unable to receive a proper number of nominations.

(*Manuscript received* 15*th September*, 1967)

ACADEMIC MOTIVATION AND SCHOOL ATTAINMENT

By N. J. ENTWISTLE

(Department of Education, University of Aberdeen)*

SUMMARY. A twenty-four item self-rating inventory was constructed to assess academic motivation. Scores on the inventory were compared with teachers' ratings to give evidence of concurrent validity. A test-retest reliability coefficient of +0·83 was obtained with an interval of 2½ months. The inventory was given to 2,707 Aberdeen 13-year-olds who had previously been given various intellectual tests. School attainment was measured from teachers' estimates scaled against a verbal reasoning test. The scores on the inventory were found to correlate more closely with school attainment than with reasoning ability. Academic motivation scores also distinguished between groups of children who had improved their academic performance after transfer and those who had deteriorated. This distinction could not be attributed simply to social class differences. Even within social class, 'improvers' had significantly higher scores on the inventory than 'deteriorators' which suggests that 'academic motivation' is not closely related to social class.

I.—INTRODUCTION.

ACADEMIC motivation is a concept which has been introduced to explain some of the differences in the school attainment of children with similar measured abilities. Differences between children in their determination to succeed in their academic studies are familiar to teachers, but assessment of these differences has proved difficult. Various attempts have been made to relate this 'academic motivation' to the more general personality trait of 'achievement motivation.' Most of the work on this 'n-ach' variable has been done in America by McClelland (1953) and Atkinson and Feather (1966), using a thematic apperception test. The results obtained when this measure has been used to predict academic achievement in school have been largely contradictory (Lavin, 1965). A more specific measure of 'academic motivation' appears to be necessary.

The most promising attempts at measuring this specific motivation have also been made in America. Schlesser and Finger (1962) devised a self-rating scale, the Personal Values Inventory, which identified attitudes and personality traits related to behaviour in school and college. In a factor analysis of the sub-scores from this inventory Finger and Schlesser (1965) showed that a factor of 'academic motivation' was clearly related to grades in high school. This factor appeared to cover 'aspiration,' 'attitudes to school' and 'study habits.' In Britain Astington (1960) used a simple unrefined scale and more recently Buxton (1966) reported the early stages of the development of two scales of 'academic achievement motivation.' In Aberdeen an 'academic motivation' scale was devised as part of a follow-up study into the age of transfer to secondary school (Nisbet and Entwistle, 1966).

II.—DEVELOPMENT OF ABERDEEN ACADEMIC MOTIVATION INVENTORY.

Following on the work of Finger and Schlesser and other similar scales derived from existing personality tests (Austrin, 1965; Taylor and Farquhar, 1966), thirty-six items were produced which appeared to cover, in the British setting, the dimensions contributing to academic motivation. The Personal Values Inventory had used both direct and indirect items to measure the children's attitudes. The indirect items asked the child to report what his teachers, parents or peer group thought of his behaviour. A similar approach was used in the pilot Aberdeen inventory.

---

* Now at Department of Educational Research, University of Lancaster.

An item analysis was used to refine this scale, based on on the responses of 150 first-year children from a Scottish comprehensive school. Three teachers were asked to put each child into one of five defined categories of motivation. Items were rejected if they failed to discriminate between groups with high or low teachers' ratings. In this way the number of items was reduced to twenty-four. Apparently indirect items such as " Do your parents think you should work harder ? " were not meaningful to Scottish 12-year-olds. The majority of such items were eliminated by this analysis. It was also found that certain items had to be modified to make the question less extreme or to avoid an indirect measurement of intelligence. The pattern of responses by the children rated highly by their teachers was used as the basis of the revised scoring scheme. A fuller description of this analysis has been given elsewhere (Entwistle, 1967).

The revised version of this inventory, shown as an Appendix, was cross-validated in another comprehensive school by comparing teachers' ratings with pupils' scores on the inventory. The results of this comparison are shown in Tables 1 and 2. The mean academic motivation scores for the various groups shown in Table 1 show significant differences between children rated high and those rated low by their teachers. But there were also significant differences in mean academic motivation score between high and low streams, suggesting that pupils with high academic motivation scores also tended to be more intelligent.

TABLE 1

Mean Scores on the Academic Motivation Inventory of Pupils in High or Low Streams Rated by their Teachers as Having either High or Low Motivation.

Streams	Rating on Academic Motivation				Significance of difference between high and low groups	
	High	(N)	Low	(N)	t	p
ABC	20·73	(41)	18·76	(67)	4·17	<·001
DE	17·50	(12)	14·68	(38)	2·24	<·05

TABLE 2

Rank-Order Correlation Coefficients between Teachers' Ratings and Scores on the Academic Motivation Inventory.

Class	N	rho	Z/σ_z	p
A	37	·38	2·33	<·05
B	36	·21	1·20	NS
C	35	·55	3·50	<·01
D	30	·04	0·02	NS
E	20	·50	2·26	<·05

Table 2 shows that all the rank-order correlation coefficients between academic motivation and teachers' ratings were positive, although only three of the values were statistically significant.

This cross-validation indicated that the scale had some concurrent validity, though the evidence was not conclusive. It was realised that the criterion of validity used was open to criticism as the pupils had only been in secondary school for a term, and thus further evidence of concurrent validity was sought.

Test-retest reliability was investigated with a group of eighty-eight children who were retested after an interval of 2½ months. A reliability coefficient of +0·81 was obtained with this sample. Correction for homogeneity raised this value to +0·83.

III.—CORRELATIONAL ANALSYIS.

As part of the follow-up study being conducted in Aberdeen, the academic motivation inventory was given to a complete age-group of children towards the end of their first year in secondary school. The analysis was limited to those children for whom complete data had been obtained on tests given between the ages of 11 and 13. This sample was about 80 per cent. of the age-group and has been shown to be representative of the age-group as a whole (Entwistle, 1967).

The mean scores on the inventory were found to be 15·55 (SD=4·54) for the boys and 16·65 (SD=3·91) for the girls. School attainment was measured using teachers' estimates averaged over five subjects, scaled against Moray House Advanced Verbal Reasoning Test 12 by the method described by Yates and Pidgeon (1957, page 88). The previous tests given to this age-group were: NFER Non-verbal Test 1 at age 11; Moray House Verbal Reasoning Test 72 at age 12; and primary teachers' estimates of attainment scaled on Moray House attainment tests at age 12. The correlations of these tests with the inventory are shown in Table 3. Allowing for the decrease in correlation as time interval increases, it is clear that the inventory correlates more highly with attainment than with reasoning ability. This result suggests that the inventory is measuring a non-intellectual trait independently related to school attainment, and gives a much more satisfactory indication of concurrent validity.

TABLE 3

Product-Moment Correlation Coefficients between Academic Motivation Inventory and Various Intellectual Measures by Sexes.

Variable	Age	Boys (N=1385)	Girls (N=1322)
Teachers' Estimate of Attainment	13	·501	·412
Verbal Reasoning	13	·385	·339
Teachers' Estimate of Attainment	12	·412	·360
Verbal Reasoning	12	·369	·326
Non-verbal Reasoning	11	·301	·283

IV.—ANALYSIS OF ' IMPROVING ' AND ' DETERIORATING ' PUPILS.

Much of the previous research on ' over-achievers ' and ' under-achievers ' has been affected by statistical regression to the mean (Thorndike, 1963). If children are identified whose attainment scores are higher or lower than their reasoning test scores, it is inevitable that the ' over-achievers ' will contain a majority of less intelligent children. These children with low reasoning test scores have more room for improvement than children with much higher scores. The differences in intelligence between such groups of over- and under-achievers prevent any valid conclusions being drawn about the improved correlates of performance.

In the present analysis of changes in attainment, regression effects were controlled by sorting into ' improvers ' and ' deteriorators ' at different levels of ability. In 1965 the children in this sample were allocated to senior or junior secondary schools on the basis of a total transfer score. This was the sum of scores on verbal reasoning tests, attainment tests and teachers' estimates of attainment, representing the best single predictor of secondary school success. Each set of test scores had been normalised to give a mean of 100 and standard deviation of 15 on this age-group. The total transfer scores were grouped into fourteen categories, each representing 5 points of quotient between 130+ and 69 –. To allow comparisons to be made between senior and junior secondary schools, the analysis was restricted to pupils with transfer scores between 119 and 100. Each category of transfer score was taken separately and sorted against the teachers' estimates of secondary school attainment at age 13. Those children whose attainment category was higher than their transfer score category formed the ' improving ' group; those with a lower attainment category were the ' deteriorating ' group. If the academic motivation scores were found to be related to these changes in attainment, further evidence of concurrent validity might be inferred.

TABLE 4

DIFFERENCES IN THE MEAN SCORES ON THE ACADEMIC MOTIVATION INVENTORY OF GROUPS OF CHILDREN WHO HAD ' IMPROVED ' OR ' DETERIORATED ' IN THEIR ACADEMIC PERFORMANCE AFTER ENTERING SECONDARY SCHOOL.

Transfer Score Category	Improving Group		Deteriorating Group		Significance of Difference	
	Mean	(N)	Mean	(N)	t	p
Senior Secondary:						
119—115	19·79	(73)	17·66	(100)	4·91	$<\cdot 001$
114—110	19·15	(89	17·25	(77)	3·39	$<\cdot 001$
109—105	19·13	(54)	16·68	(22)	2·27	$<\cdot 05$
Junior Secondary:						
114—110	16·90	(15)	16·61	(35)	0·26	NS
109—105	18·50	(85)	15·49	(124)	5·47	$<\cdot 001$
104—100	17·12	(113)	14·81	(147)	4·27	$<\cdot 001$

The mean scores of the ' improving ' and ' deteriorating ' groups for each of the four transfer categories considered are shown in Table 4. These results show that in all but one of the ability groups the mean academic motivation score of the ' improving ' group was significantly higher than that of the 'deteriorating' group. Combining the transfer score categories in the restricted range of ability chosen does not alter this pattern. Table 5 shows the differences between transfer scores and secondary school attainment scores for the two groups and shows highly significant differences in mean academic motivation scores between the groups.

TABLE 5

DIFFERENCES IN THE MEAN SCORES OF COMBINED GROUPS OF 'IMPROVERS' AND 'DETERIORATORS' IN SENIOR OR JUNIOR SECONDARY SCHOOLS.

Measure	Improving Group		Deteriorating Group		Significance of Difference	
	Mean	(N)	Mean	(N)	t	p
Senior Secondary:						
Secondary School Attainment Estimate	119·94	(216)	106·30	(199)	—	—
Transfer Total	112·44	—	113·96	—	—	—
Academic Motivation ..	19·36	—	17·39	—	5·99	<·001
Junior Secondary:						
Secondary School Attainment Estimate	111·93	(213)	98·00	(306)	—	—
Transfer Total	104·70	—	105·17	—	—	—
Academic Motivation ..	17·66	—	15·28	—	6·37	<·001

The possible social origins of 'achievement motivation' have recently been discussed by Swift (1966). It was thus anticipated that the scale of academic motivation might be an indirect measure of social class. The analysis was thus repeated within social class. For this purpose three bands of father's occupation were formed on the basis of Fraser's (1959) previous findings in Aberdeen. These bands were: (*a*) social classes I and II on the Registrar General's Scale; (*b*) social class III (non-manual and highest skilled manual) and (*c*) social classes III (remaining occupations), IV and V. The mean scores of 'improving' and 'deteriorating' groups on the academic motivation inventory, within social class, are shown in Table 6. These results show that the academic motivation

TABLE 6

DIFFERENCES IN THE MEAN SCORES ON THE ACADEMIC MOTIVATION INVENTORY OF 'IMPROVERS' AND 'DETERIORATORS' WITHIN SOCIAL CLASS.

Social Class	Type of Secondary School	Improving Group		Deteriorating Group		Significance of Difference	
		Mean	(N)	Mean	(N)	t	p
High (I, II)	Senior	19·43	(82)	17·33	(41)	2·83	<·01
	Junior	17·41	(22)	15·68	(34)	1·46	NS
Intermediate (III a, b)	Senior	19·88	(74)	17·43	(83)	4·88	<·001
	Junior	17·81	(98)	15·73	(101)	3·54	<·001
Low (IIIc, IV, V)	Senior	18·35	(40)	17·23	(66)	1·77	NS
	Junior	17·38	(80)	14·91	(153)	4·28	<·001

scores distinguish between 'improvers' and 'deteriorators' within each social class, although the differences are not significant in every group. Certainly academic motivation scores are related to social class, but not to any great extent. This dimension describes marked differences in children's behaviour within the social classes and these differences are related to changes in school attainment.

V.—DISCUSSION OF RESULTS.

The implications of the present findings seem to be that social class is not closely related to 'academic motivation' as measured by the Aberdeen Inventory. The theory of the origins of this motivation has not been developed far enough to give satisfactory explanations. Certainly a child's achievement motivation is likely to reflect parental attitudes and patterns of child rearing (Rosen and d'Andrade, 1959). But the Plowden Social Survey (1967) has shown that there are wide differences in 'parental aspiration' within social class, and that this measure shows a much greater independent relationship with primary school attainment in reading than does the variable of 'father's occupation.' Swift (1967) has stressed that 'social class' is too broad a measure to describe the different home environments experienced by children. The significant differences in academic motivation within social class found in the present investigation reinforce this conclusion. But this may not be a complete explanation. We cannot assume that all children react in the same way to similar environments. Eysenck (1957) has suggested that inherited differences in 'conditionability' may affect children's reactions to their environment. The attitudes and behaviour making up 'academic motivation' presumably represent the result of interactions between the child's environment and his basic temperament, rather than simply more specific environmental variables within social class. It seems likely that any attempt to formulate a satisfactory theory of 'academic motivation' must take account of both temperamental traits and environmental variables.

The present study, however, was not an attempt to construct such a theory: its aim was to measure academic motivation by a self-rating inventory. The design of such an inventory demands caution. The questions must be meaningful to the child and must also avoid any undesirable mental set. Most indirect items were found to be unsuitable for Scottish children and the other questions were chosen to avoid any implied value judgments on the behaviour being reported. The results obtained from the correlational analysis and from the analysis of 'improvers' and 'deteriorators' show that this scale is measuring a behaviour trait, other than intelligence, which is related to school attainment. The content of this inventory covers attitudes to school, aspirations, work habits and the determination to do well in school work. It seems reasonable then to assume that this scale has been effective in measuring 'academic motivation.'

ACKNOWLEDGEMENTS.—This project has been supported by a grant from the Scottish Council for Research in Education. Head teachers and teachers in various parts of Scotland co-operated in the development of the inventory. In Aberdeen town and county, many teachers gave invaluable help by supplying detailed orders of merit for their classes and by administering the inventory. All this assistance is gratefully acknowledged by the author.

VI.—References.

Astington, E. (1960). Personality assessments and academic performance in a boys' grammar school. *Brit. J. Educ. Psychol.*, 30, 225-236.

Atkinson, J. W., and Feather, N. T. (1966). *A Theory of Achievement Motivation.* New York : Wiley.

Austrin, H. R. (1965). Cross-validation of an attitude scale for the identification of high and low academic achievers. *J. Educ. Res.*, 58, 9, 426-428.

Buxton, C. E. (1966). Evaluations of forced-choice and Likert-type tests of motivation to academic achievement. *Brit. J. Educ. Psychol.*, 36, 192-201.

Entwistle, N. J. (1967). *The Transition to Secondary Education : the Age of Transfer and Correlates of Academic Success in the First Year at Secondary School.* Ph.D. thesis, Aberdeen University.

Eysenck, H. J. (1957). *Sense and Nonsense in Psychology.* London : Penguin Books.

Finger, J. A., and Schlesser, G. E. (1965). Non-intellective predictors of academic success in school and college. *School Rev.*, 73, 14-29.

Fraser, E. (1959). *Home Environment and the School.* London : Univ. of London Press.

Lavin, D. E. (1965). *The Prediction of Academic Success.* New York : Russell Sage Foundation.

McClelland, D. C. (1953). *The Achievement Motive.* New York : Appleton, Century, Crofts.

Nisbet, J. D., and Entwistle, N. J. (1966). *The Age of Transfer to Secondary Education.* London : Univ. of London Press.

Plowden Social Survey (1967) in *Children and their Primary Schools* (Volume 2). London : H.M.S.O.

Rosen, B. C., and D'Andrade, R. (1959). The psycho-social origins of achievement motivation. *Sociometry*, 22, 185-218.

Schlesser, G. E., and Finger, J. A. (1962). *Personal Values Inventory.* Hamilton, New York : Colgate Univ. Press.

Swift, D. F. (1966). Social class and achievement motivation. *Educ. Res.*, 8, 83-95.

Swift, D. F. (1967). Family environment and 11+ success : some basic predictors. *Brit. J. Educ. Psychol.*, 37, 10-21.

Taylor, R. G., and Farquhar, W. W. (1966). The validity and reliability of the human trait inventory designed to measure under-and over-achievement. *J. Educ. Res.*, 59, 227-230.

Thorndike, R. L. (1963). *The Concepts of Over- and Under-achievement.* New York : Teachers' College, Columbia University.

Yates, A., and Pidgeon, D. A. (1957). *Admission to Grammar Schools.* London : Newnes and N.F.E.R.

VII.—Appendix.

ABERDEEN ACADEMIC MOTIVATION INVENTORY

Here are some questions about school and school work. In all the questions you *must* answer either 'yes' or 'no.' Put a circle round the answer you wish to give. Answer ALL the questions TRUTHFULLY but quickly.

		'Correct' Response
1.	Do you like being asked questions in class ?	Yes.
2.	Does your mind often wander off the subject during lessons ?	No.
3.	Do you enjoy most lessons ?	Yes.
4.	Do your parents want you to start work when you are 15 ?	No.
5.	Do you think school is rather a waste of time ?	No.
6.	Do you like to leave your homework to the last minute ?	No.
7.	If you were given lower marks than usual in a test, would this make you unhappy ?	Yes.
8.	Do you expect school to provide you with good qualifications for a job ?	Yes
9.	Is it important to you to do well at school ?	Yes.
10.	Are you happier working with your hands ?	No.
11.	When you are given a difficult problem, do you enjoy trying to find the answer ?	Yes.

12. Do your parents expect you to go to university or college ? Yes.
13. Do you generally find lessons rather dull ? No.
14. Do you dread being given a test on your homework ? No.
15. Do your friends think that you never take work seriously ? No.
16. Would you like to leave school as soon as possible ? No.
17. Do your parents tell you to enjoy yourself and not to worry about school ? No.
18. Do you work hard most of the time ? Yes.
19. Do your parents think that you must do well at school if you are to succeed in later life ? Yes.
20. Do your teachers think that you misbehave too much ? No.
21. Do you worry about not doing well in class ? Yes.
22. Are you more interested in games than school work ? No.
23. Do you find it difficult to keep your mind on your work ? No.
24. Do you always try your hardest to get your homework right ? Yes.

(*Manuscript received* 15*th October*, 1967)

OPERATIONAL CONSERVATION OF ILLUSION-DISTORTED LENGTH

By FRANK B. MURRAY

(*Department of Educational Psychology, University of Minnesota*)

SUMMARY. Two sticks of equal length were made to look unequal by the Muller-Lyer illusion. A child conserved length if he saw the illusion and despite it, maintained that the lengths of the sticks were the same. Subjects were directed to use one of the sticks that looked unequal in the Muller-Lyer illusion and then if they wanted to, the other stick, to fish a coin from a cage. A child failed to conserve the length of the sticks if he refused or used the second stick and said he did because it was longer or shorter than the first stick. A significant number of children ($p < \cdot 001$) who conserved length on the first task did not conserve it on the second.

I.—INTRODUCTION.

A CHILD is considered to conserve the length of a thing when he recognizes that the length of a thing can change only by the operations of addition and subtraction, or in some cases by expansion or compression. Should he judge that the length of a thing changes when it is subjected to such irrelevant transformations as mere position change or perceptual distortion, the child is considered not to conserve length. Whatever concept of length the non-conserving child uses, it is apparently defective insofar as he claims the length of a thing has changed when it in fact has not.

In the common test for conservation of length, children are shown two things of equal length which are subsequently transformed so that their lengths appear unequal. The child is then asked something like, " Are both things the same length or is one longer or shorter " ? The question may be ambiguous to the young child who may interpret " Which is longer ? " to mean " Which looks longer ? " The child who makes this error will, of course, be scored as a non-conserver when had some other less ambiguous measure of conservation been used the child might have indicated that he was aware that the objects in question were the same length. In a previous study (Murrary, 1965) the child was asked both questions, namely, " Which is . . . ? " and " Which looks . . . ?," to guard against the possible misinterpretation. Nevertheless, the possibility of misinterpretation, though diminished, still existed and it was suggested that conservation be assessed by something more than the child's responses to these questions.

Braine (1959) has examined the transitivity of length, for example, with a non-verbal method and found that the concept was acquired two years earlier than Piaget's age norms ; Braine's method has been criticized (Smedslund, 1963) and the controversy has continued (Braine, 1964, and Smedslund, 1965). In the present experiment a task was devised which eliminated the possibly ambiguous questions. This task required the subjects to act upon, or operate with, the equality they had conserved verbally ; subjects who did so were said to be operational length conservers. It was the difference between length conservation, as measured by the children's responses to the two questions, and operational length conservation that was analyzed in this study.

II.—METHOD.

Subjects.

All 119 subjects, sixty-nine boys and fifty girls, were enrolled in the kindergarten, first and second grades of the Lida Lee Tall Laboratory School in Baltimore, Maryland. The mean age for each grade was 5·76 years (SD=·30, range 5·25—6·16), 6·71 years (SD=·33, range 6·25—7·16), and 7·79 years (SD=·34, range 7·25—8·16), respectively. The mean IQ from the SRA Primary Mental Abilities Test was 112·4 (SD=11·4) for the first grade and 112·5 (SD=10·9) for the second grade. Test results for the kindergarten were not available. There were in all thirty-one kindergartners, forty-two first-, and forty-six second-graders.

All testing was done individually in a small conference room in the laboratory school. Each subject faced the experimenter across a table which held the stimulus materials used in the study.

Materials.

The materials for the length conservation test were: (1) ten round white sticks, 5 mm. in diameter, two of which were the same length (21 cm.), four were unequally longer than 21 cm. and the rest, shorter, (2) a black rectangular (30 cm.×50 cm.) composition board on which were glued three cardboard figures—the arrows and feathers of the Müller-Lyer figure in horizontal position. The two segments of the arrows and feathers were 5 mm.×7 cm. and at right angles to one another.

The additional materials for the test of operational length conservation were a 25 cent piece and a 29 cm.×20·5 cm.×18 cm. stainless steel wire cage.

Procedure.

The experimenter introduced all tasks as a game, and explained his note-taking as keeping score so he would not forget how well the subjects did. Each subject was asked if he had any idea of what he was about to do.

Before the test of length conservation, subjects were presented a long and a short stick from the ten sticks described above, and asked to find the longer one. Subjects were asked then if they could find two sticks in the ten that were the same length or size. If the subject had any difficulty with the multiple discriminations, the experimenter prompted by pointing to two unequal sticks and asking about their length, and then pointing to the two equal sticks and asking about their length. The children's responses to these preliminary questions ensured that the subjects understood the directions in the way that the experimenter intended.

The experimenter took the equal sticks that the subject had selected and placed them uprightly side by side on the table to enable the subject to reconfirm their equality. Then the experimenter slowly placed them in the Muller-Lyer configuration to make their lengths appear unequal. One-half the subjects in each grade had the 'feathered figure' on the left, and the rest had it on the right. Subjects were asked:

> Does this stick (experimenter pointed to the stick in the 'feathered' configuration) look longer than this stick (experimenter pointed to the 'arrowed' stick), does it look the same length as this stick, or does it look shorter than this stick?

To nullify any leading influence the order of the alternatives (longer, shorter, the same length as) might have on the subject's response, the order of the alternatives was varied randomly among the subjects. If a subject answered that the sticks looked the same length, i.e., he failed to see the illusion, he was discarded from the experiment.

Subjects were asked a second question:

If the sticks were standing up the way they were before, would this stick (experimenter pointed to the stick in the feathered configuration) be longer than this stick (experimenter pointed to the other stick), would it

---

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOTICES

Subscriptions and Orders.

As from January, 1967, subscriptions and orders for this Journal should be sent to: The Manager, B.J.E.P. Department, Sweet and Maxwell, Spon Booksellers, Ltd., 11, New Fetter Lane, London, E.C.4.

Papers for publication and books for review, should, as previously, be sent to the Editor at The Institute of Education, The University, Sheffield, 10.

Reprinting of February Number of Volume 36.

It is very much regretted that, due to unpredicted heavy demands, the February Number of Volume 36 of this Journal is out of print and a number of subscribers have not received copies. Reprinting is taking place and copies will be available shortly.

L. B. BIRCH,
Editor.

II.—METHOD.

Subjects.

All 119 subjects, sixty-nine boys and fifty girls, were enrolled in the kindergarten, first and second grades of the Lida Lee Tall Laboratory School in Baltimore, Maryland. The mean age for each grade was 5·76 years (SD=·30, range 5·25—6·16), 6·71 years (SD=·33, range 6·25—7·16), and 7·79 years (SD=·34, range 7·25—8·16), respectively. The mean IQ from the SRA Primary Mental Abilities Test was 112·4 (SD=11·4) for the first grade and 112·5 (SD=10·9) for the second grade. Test results for the kindergarten were not available. There

To nullify any leading influence the order of the alternatives (longer, shorter, the same length as) might have on the subject's response, the order of the alternatives was varied randomly among the subjects. If a subject answered that the sticks looked the same length, i.e., he failed to see the illusion, he was discarded from the experiment.

Subjects were asked a second question:

> If the sticks were standing up the way they were before, would this stick (experimenter pointed to the stick in the feathered configuration) be longer than this stick (experimenter pointed to the other stick), would it be the same length as this stick, or would it be shorter than this stick?

Again, the order of the alternatives was varied randomly among the subjects.

The two sticks were shown again to be equal by standing them next to one another before the subject. The sticks were placed in the Müller-Lyer figure once more so that one looked longer than the other. Subjects were directed to pick up one of the sticks from the Muller-Lyer figure and attempt by poking it through the wire cage mesh to reach a quarter and move it to the front of the cage. The coin was placed, however, so that it could be reached, but not moved closer to the subject. After subjects vainly attempted for thirty seconds to move the coin. the experimenter asked each subject if he thought he could do better at the task with the other stick. If the subject picked up the second stick, he was asked by the experimenter why he thought he could do better with it.

III.—RESULTS.

All subjects answered the preliminary questions about unequal and equal length correctly. Each subject on the first task was scored as having conserved length only if (1) he saw the Müller-Lyer illusion, that is, said the feathered stick looked longer, and (2) despite seeing the illusion maintained in response to the second question that the sticks were the same length. If a subject maintained that the sticks were unequal in length, he was scored, of course, as a nonconserver.

A subject was scored as a conserver in the operational length conservation task if he refused the second stick and said he did so because both sticks were the same length; he was scored a nonconserver, however, if he refused the second stick and said he did so because it was shorter. A subject was scored a nonconserver also, if he attempted to use the second stick and said he did so because it was different from the first stick in length. If the subject could give no reason for wanting to use the second stick or if he said that he simply wanted to try it, nothing could be inferred about length, conservation or nonconservation and the subject was discarded from the analysis. There were ninety-six subjects who could be scored as conservers or nonconservers in the task.

The comparison between the performance of these subjects on the length conservation task and the operational length conservation task showed a strong shift from length conservation to operational length nonconservation (Table 1). The McNemar Test (Siegel, 1956) was applied to the data in table 1 and indicated that the shift was significant ($\chi^2=15\cdot42$, $p<\cdot001$).

The group was divided at its median age (7·08 years) into an older and younger group. The McNemar Test, applied to the older and younger groups' scores (Table 1), indicated that the shift from length conservation to operational length nonconservation was significant for both groups, but more so for the younger group ($\chi^2=9\cdot09$, $p<\cdot01$) than for the older group ($\chi^2=4\cdot90$, $p<\cdot05$). The percentage of subjects who shifted from conservation to non-conservation for age groups five years ($N=13$), six years ($N=31$), seven years ($N=35$), and eight years ($N=17$) was 7·6 per cent, 32 per cent, 17 per cent and 12 per cent,

TABLE 1

NUMBER OF SUBJECTS WHO CONSERVED LENGTH ON THE LENGTH CONSERVATION AND OPERATIONAL LENGTH CONSERVATION TASKS FOR THE TOTAL, OLDER AND YOUNGER-SUBJECT GROUPS.

Length Conservation	Operational Length Conservation	
	Nonconservers	Conservers
Total (N=96) :		
Conservers	20	45
Nonconservers	30	1
Younger (N=48) :		
Conservers	11	14
Nonconservers	23	0
Older (N=48) :		
Conservers	9	31
Nonconservers	7	1

respectively. The median test (Siegel, 1956) applied to the data indicated a significant difference in operational length conservation for the group above the median age (7·08 years) and below it ($\chi^2=12{\cdot}06$, $p<{\cdot}001$).

IV.—DISCUSSION.

The shift from conservation to operational nonconservation may have resulted from nothing more than a mere measurement artifact stemming from a looseness in defining conservation, as Gruen (1966) has suggested in his observation that, when a more demanding criterion is used, fewer young subjects can be scored as conservers. However, when one considers the nature of the tasks and that both tasks required judgments about the same sticks from the same children, the finding seems to go deeper than a measurement artifact. The finding of a shift from conservation to nonconservation is consistent with the Genevan observations in the child's development of horizontal *décalages* (Szeminska, 1965), namely, the observation that the same child can exhibit operational stage thinking in one task and yet lag in exhibiting operational thinking and exhibit pre-operational thinking on tasks of more complex content.

Although the mechanisms of horizontal *décalages* are not well understood, a number of factors may be hypothesized to have caused subjects who had conserved on the first problem to switch to non-conservation on the coin and cage problem. Perhaps the most attractive of these is the notion of an alternating dominance between older perceptually bound thinking and newer logical thinking. Halpern (1965) has show that, even in the presence of logical structures, perception can govern thinking. The progression from one mode of thinking to another is clearly not an all-or-nothing development ; the perceptual and logical strategies are not mutually exclusive.

It is well accepted that under particular conditions, e.g., stress, organisms resort to lower levels of functioning even when higher levels are available. The operational coin and cage conservation problem, it might be argued, subjected the children to more stress since it was clear to them that their performance on the task could be right or wrong, i.e., they moved the coin or they didn't. During the first task subjects were not under this stress, and the reduction in risk favoured the newer, less developed strategy. The interpretation of the data as a regression to lower order operations is supported further by the finding that the conservation-nonconservation shift was stronger among the younger subjects than it was among the older subjects.

Besides perhaps differing in stress, the length conservation problem and operational length conservation problem differed also in that the latter did not use the questions about appearance and reality. Braine and Shanks (1965) contend that children under seven tend to interpret the question "Which is longer?" in terms of phenomenal rather than real size. Consequently, they are scored as nonconservers when they might not be. This approach does not explain nonconservation, but merely transfers the problem to a consideration of linguistic development; one still has to explain why conservers understand the questions while nonconservers do not. The present data, in any case, do not agree with Braine (1959) and Braine and Shanks (1965) that when the ambiguous verbal questions are eliminated, conservation will be manifested at a younger age.

A parsimonious explanation may be that what was operating in the shift to nonconservation was the well known difference, or lag, between recognition and reproduction. In another context, Maccoby and Lee (1965) have referred to this as the lag between perceiving and performing, namely, that children can make discriminations in stimuli before they can reproduce them. In the present study, young children may recognize conservation before they can reproduce conservation or act upon it.

ACKNOWLEDGEMENTS.—This experiment was part of a doctoral dissertation at the Johns Hopkins University. The co-operation of the faculty and students of the Lida Lee Tall Laboratory School in Baltimore, Maryland, was appreciated.

V.—REFERENCES.

BRAINE, M. (1964). The development of a grasp of transitivity of length: a reply to Smedslund. *Child Developm.*, 35, 799-810.

BRAINE, M. (1959). The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychol. Monogr.*, 73, 5 (Whole No. 475).

BRAINE, M., and SHANKS, BETTY (1965). The development of the conservation of size. *J. Verb. Learn. Verb. Behav.*, 4, 227-242.

GRUEN, G. (1966). Note on conservation: methodological and definitional considerations. *Child Developm.*, 37, 977-983.

HALPERN, ESTHER (1965). The effects of incompatibility between perception and logic in Piaget's stage of concrete operations. *Child Developm.*, 36, 491-498.

MACCOBY, ELEANOR and LEE, HELEN (1965). Some speculations concerning the lag between perceiving and performing. *Child Developm.*, 36, 376-378.

MURRAY, F. (1965). Conservation of illusion distorted lengths and areas by primary school children. *J. Educ. Psychol.*, 56, 62-66.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

SMEDSLUND, J. (1963). Development of concrete transitivity of length in children. *Child Developm.*, 34, 389-405.

SMEDSLUND, J. (1965). The developmental transitivity of length: a comment on Braine's reply. *Child Developm.*, 36, 577-580.

SZEMINSKA, ALINA (1965). The evolution of thought: some applications of research findings to educational practice. In Paul Mussen (Ed.), European research in cognitive development. *Monogr. Soc. Res. Child Developm.*, 35, 100, 47-57.

(*Manuscript received 20th October, 1967*)

RESEARCH NOTES

THE CORRELATION OF GCE A-LEVEL GRADES WITH UNIVERSITY EXAMINATIONS IN CHEMICAL ENGINEERING

By DOUGLAS G. BAGG

(*University of Manchester Institute of Science and Technology*)

SUMMARY. Secondary school A-level science examination grades of university students of chemical engineering were compared with first-year and final university examination marks. Three groups were studied: 556 successful U.K. students, fifty-six successful overseas students and sixty-seven unsuccessful students. In the first group, the A-level Physics mark showed the highest correlation with first-year and final performance, and Mathematics the lowest. For overseas students, the position of these two subjects in the correlations was reversed.

I.—INTRODUCTION.

Research is in progress into the relations between various factors in the academic and biographic records of several hundred university students of chemical engineering. This is a preliminary report of some of the findings. Places are offered to students partly on the basis of their A-level grades and it is therefore important to know how closely these correlate with subsequent academic performance. Previous research has indicated that the level of correlation to be expected is quite low (Drever, 1963) and also that the value obtained is likely to vary from subject to subject (Austwick, 1960). This latter finding was recently confirmed by the author in the analysis of a small sample of records from a technology other than chemical engineering, when the correlation coefficients with all three of the science A-levels were found to be negative. Richards and Wilson (1961) reported that the failure rate of students entering with the highest A-level grade was as much as 15 per cent. The present study has so far investigated the correlation between A-level science grades and the marks obtained in the first-year sessional examinations or in the Part II Finals (third year).

II.—THE PRESENT STUDY.

Results are reported for three groups of students from a population of 807 admissions during the years 1954-65, inclusive. The main sample consisted of 556 U.K. students who had completed their degree course by June 1967, and a similar analysis was carried out on a group of fifty-six overseas students (from the same population of 807) who completed the course, having first taken A-levels set by English Boards. A subsidiary analysis was performed on a group of sixty-seven U.K. students from this population who had failed to complete the course. The remaining records were incomplete and have been omitted from this preliminary report.

It was found that the frequency distribution of grades obtained in the A-level examinations (A—E) was skewed, 30–40 per cent of the entrants having grade B passes, and the grades for each subject were therefore converted into normalised T-scores with a mean of 50 and a standard deviation of 10. To form another predictor, the A-levels were summed for each student, using a 5–4–3–2–1 scale in which grade A=5, and this figure was converted similarly to a normalised T-score. The frequency distributions of the criteria, namely the first year sessional and the Part II Finals examinations, were not seriously skewed and these criteria were used as raw scores throughout.

III.—RESULTS FROM BRITISH STUDENTS COMPLETING THE COURSE.

The means and standard deviations of this sample of 556 students are as follows:

	A-levels			Summed	University	
	Chemistry	Physics	Mathematics	A-levels	1st Year	Finals
Mean........	50·01	50·47	49·98	50·23	53·87	58·92
S.D.	8·69	9·18	9·44	9·76	9·60	7·82

The zero-order correlation coefficients were calculated by the Pearson product-moment formula and are as follows:

1st Year Sessionals and—	
Chemistry	+0·377
Physics	+0·449
Mathematics	+0·321
All—summed	+0·521

Part II Finals and—	
Chemistry	+0·260
Physics	+0·325
Mathematics	+0·195
All—summed	+0·357
1st Year Sessionals	+0·491

The intercorrelation coefficients between the A-level subjects were also calculated. Chemistry was found to correlate +0·420 with Physics and +0·176 with Mathematics. The correlation between Physics and Mathematics was +0·254.

The following regression equation was calculated for the prediction of the Part II Finals mark from the grades of GCE A-level pass expressed as normalised T-scores:

$$\text{Finals mark} = 37{\cdot}81 + (0{\cdot}1270 \times \text{chem}) + (0{\cdot}2028 \times \text{phys}) + (0{\cdot}0905 \times \text{maths}).$$

The multiple correlation coefficient, using these three A-level subjects, was found to be +0·368. The standard error of prediction was 7·27 marks. This indicates that slightly over one-third of the obtained marks for the Part II Finals should be within about 7·3 points of the values predicted by calculation from the above regression equation (Guilford, 1956, page 400). The value of the multiple correlation coefficient indicates that only 13·5 per cent of the variance of the Part II Finals marks can be accounted for by the three A-level results taken together. The remaining 86·5 per cent of this variance must be accounted for by factors unrelated to these GCE results.

IV.—RESULTS FROM OVERSEAS STUDENTS COMPLETING THE COURSE.

The means and standard deviations of the sample of fifty-six overseas students, who completed the same course as the U.K. students, are as follows:

	A-levels			Summed	University	
	Chemistry	Physics	Mathematics	A-levels	1st Year	Finals
Mean........	47·79	49·57	51·34	49·38	54·81	57·79
S.D.	8·56	9·64	10·24	9·47	8·80	8·21

Their zero-order correlation coefficients are as follows:

1st Year Sessionals and—	
Chemistry	+0·365
Physics	+0·257
Mathematics	+0·404
All-summed	+0·528

Part II Finals and—	
Chemistry	+0·289
Physics	+0·135
Mathematics	+0·229
All-summed	+0·339
1st Year Sessionals	+0·675

The intercorrelation coefficients between the A-level subjects in this rather small sample were again calculated. Chemistry correlated +0·165 with Physics and +0·112 with Mathematics. The correlation between Physics and Mathematics was +0·208.

V.—RESULTS FROM U.K. STUDENTS WHO FAILED TO COMPLETE THE COURSE.

The unsuccessful students in the population of 807 who took the first-year sessionals but failed to complete the course make a sample of sixty-seven U.K. students. The means and standard deviations of this sample are as follows:

	A-levels				University	
	Chemistry	Physics	Mathematics	Summed A-levels	1st Year	Finals
Mean........	44·66	45·46	47·55	44·00	34·24	—
S.D.	8·43	8·13	10·86	9·59	9·58	—

Their zero-order correlation coefficients are:

1st Year Sessionals and—

Chemistry	+0·364
Physics	+0·203
Mathematics	+0·114
All-summed	+0·312

In this sample the Chemistry A-level score correlated +0·342 with Physics and +0·156 with Mathematics. Physics correlated +0·150 with Mathematics.

The standard deviations in the A-level grades of these unsuccessful students are sufficiently similar to those for the successful group to indicate that homogeneity is unlikely to have appreciably affected the correlation coefficients.

VI.—DISCUSSION OF RESULTS.

For the successful students, Mathematics appears to be the least, and Physics the most, important of the three A-level subjects. The latter is understandable, since a great deal of applied physics appears in several sections of the course such as thermodynamics and the flow of fluids. Since this application generally requires considerable use of mathematics, the position of this subject as third in order of importance is a little surprising and not readily explicable.

When the results from overseas students are considered, perhaps the most surprising point is the interchange in position of mathematics and physics in their correlation with first-year sessionals, chemistry remaining just about the same as for U.K. students. Then again, the inter-correlations between the predictors are different, possibly due as much as anything to the small size of the sample. The much larger correlation (+0·675) of the first-year sessionals with the Finals (together with other larger correlations not reported here) suggests that the overseas students tend to work more steadily throughout the course than do the U.K. students.

So far as the successful students are concerned, both overseas and U.K., the size of the correlation coefficients for the summed A-levels suggests that this might perhaps be a safer predictor than the individual results, irrespective of the distribution of the latter.

For the unsuccessful students, Mathematics again shows the lowest correlation, but here Chemistry gives the highest coefficient, instead of Physics. Their mean A-level results were only 5–10 per cent below those of the successful students and the standard deviation was, if anything, slightly greater. Their first-year sessional mark was, however, less than two-thirds of that of the successful people, although the standard deviations of this mark were almost identical.

The formation of any conclusions from the results of the unsuccessful students, quite apart from its being a small sample, is rather difficult because the reason for their quitting is generally not known. Two main reasons are presumed to exist, namely, the loss of interest caused by failure at an examination, and financial difficulties consequent upon the lengthening of the course by a year. Since the A-level performances of the sixty-seven unsuccessful students actually varied from EEE to BBB, with one even at AAA, for all we know, a proportion of them might have obtained good ordinary degrees had they re-taken the examination in which they failed. In fact, judging only from these A-level totals, the author would have expected about forty-three honours degrees from these sixty-seven entrants, had the behaviour of the main sample of 556 students been followed.

For the attempted prediction of university performance from A-level results, Drever (1963) suggests that "within certain subjects we can predict first-year university performance quite well," and it is noticeable that, in chemical engineering, the correlation coefficients for the first-year results with A-levels are markedly greater than those for Part II Finals, confirming Drever's suggestion.

For comparison with the findings of Richards and Wilson (1961), the percentage of students achieving honours degrees was calculated for different A-level totals. The percentage rose from 59–61 per cent for those entering with A-level totals of 5, 6 or 7, to 70–75 per cent for those with 8, 9 or 10, 80–85 per cent for those with 11 or 12, and reached 90–95 per cent for those with a total of 13 or 14. For entrants with a total of 15, or three grade A passes, the percentage of honours degrees drops to barely 82 per cent, but this is derived from a group of only sixteen students.

Tables have been prepared giving the probabilities, at the different grades of pass, that a student will attain a degree in chemical engineering for various A-level marks in mathematics, physics, and chemistry. Codies are available from the author.

Acknowledgements.—The author is grateful to Professor Frank Morton for his permission to use the student records for this work, and to Professor F. W. Warburton for his guidance from the aspect of educational psychology.

VII.—References.

Austwick, K. (1960). G.C.E. to B.A. *Univ. Quart.*, 15, 64-71.

Drever, J. (1963). *Prediction, Placement and Choice in University Selection*. Godfrey Thomson Lecture. Edinburgh: Moray House.

Guilford, J. P. (1956). *Fundamental Statistics in Education and Psychology*. New York: McGraw Hill.

Richards, J. P. G., and Wilson, A. J. C. (1961). 'A'-level and pass degree in Physics. *Univ. Quart.*, 15, 389-391.

(*Manuscript received 5th August*, 1967)

A NOTE ON THE WALLACH-KOGAN TESTS OF CREATIVITY

By A. J. CROPLEY

(*University of Saskatchewan Regina Campus, Canada*)

Summary. The Wallach-Kogan tests of creativity were administered to 124 first-year university men under group conditions, along with five tests of intelligence. Correlation coefficients indicated that the battery of creativity tests possessed a high degree of internal consistency, and relative independence of intelligence tests. However, a substantial general factor loaded highly on both kinds of test.

I.—Procedure.

The Wallach-Kogan tests of 'creativity' (Wallach and Kogan, 1965) were administered to 124 first-year university men, newly enrolled at the University of New England, Australia, in March, 1967. The tests were administered during a larger three-day testing project in which all first-year students participated as part of their enrolment procedures. The men, whose mean age was 18·4 years, comprised just over half the total new male enrolment for 1967. As attendance at the testing sessions was voluntary, many students did not complete the entire battery.

In keeping with the Wallach-Kogan specifications for administration of the tests, an attempt was made to create an informal atmosphere during administration.

No time limits were imposed, and students were told that they could work for as long or as short a time as they wished. They were also told that the tests were described as games, when being used with younger children, and that they were at liberty to think of them in that way. However, an important departure was made from the Wallach-Kogan procedures in that the tests were administered in a group form, and subjects had to write down their own answers rather than merely say them and have them recorded by the experimenter. The range of individual total times for the five tests was from about half-an-hour to about three-and-one-half hours.

Also administered during the admission testing procedures were five tests of the more conventional kind, tests characterized by having single ' correct ' answers which students were required to find. These five tests of ' intelligence ' included the AL test of verbal intelligence, the AQ test of numerical intelligence, and the Vocabulary, Speed, and Level of Comprehension subtests of the Co-operative Reading Tests. All five of these tests are published by the Australian Council for Educational Research (ACER), and are specifically designed for use with students at senior high school and university level. Thus, they avoid to some extent the restriction of variance likely to be present in ' intelligence ' scores obtained by any highly selected sub-group like first-year university students.

Scoring of the conventional tests was carried out according to the published directions, while the creativity tests were scored for fluency (sheer number of responses) on the one hand, and uniqueness (number of responses not found in the protocols of any other member of the group of 124 men) on the other. This is the procedure recommended by Wallach and Kogan (1965, pp. 30-31). Raw AL and AQ scores were not converted to IQs, since this procedure involves corrections for age, whereas other variables are not age-corrected. Scoring each of the five creativity tests in two ways yielded ten sets of scores so that, with the five convergent tests, scores on fifteen variables were available. Product-moment correlations among these fifteen variables were then calculated.

II.—Results.

As correlational procedures were employed in the study, it is worth noting that very wide ranges of scores were obtained on all variables. The AL scores ranged from 5 to 26, and the AQ scores from 2 to 24, so that virtually the total range of possible scores was obtained by the men involved in the study. Ranges for the other three intelligence tests were also very wide, including 0 to 54 for Level of Comprehension, 2 to 55 for Speed of Comprehension, and 9 to 58 for Vocabulary. (These are not raw scores, in the case of the two comprehension tests, but are transformed scores obtained according to norms based on highly selected groups like the university men involved in the present study.) Thus, small correlations between creativity and intelligence tests cannot be dismissed as an artifact of reduced variance in the intelligence variables.

Intercorrelations among the creativity tests scored for fluency, the same tests used for uniqueness, and the five intelligence tests, respectively, are shown in Table 1.

III.—Discussion.

Wallach and Kogan (1965, p. 45) have reported correlations among their tests big enough to warrant the conclusion that the battery of tests implies a pervasive and generalized cognitive mode appropriately labelled ' creativity.' Furthermore, cross-correlations between their creativity tests and intelligence tests were sufficiently low to support the view that the cognitive mode thus indicated was substantially independent of that elicited by tests of the conventional kind.

The data reported here lend modified support to their conclusions. It is apparent that there is a substantial degree of internal consistency within the creativity batteries, as almost all within-battery correlations are significant beyond the ·01 level of confidence. At the same time, none of the cross-correlations among creativity and intelligence measures reaches this level of significance. In other words, although as much as 55 per cent of the variance of Line Meanings (fluency),

TABLE 1

MEANS, STANDARD DEVIATIONS AND INTERCORRELATIONS FOR ALL VARIABLES.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean	45·4	34·3	20·9	19·8	20·4	2·26	2·15	1·59	4·89	4·20	17·0	14·8	39·0	27·4	20·3
S.D.	19·6	14·5	12·2	8·70	9·47	2·85	2·51	2·33	3·43	3·39	4·48	4·64	10·6	12·3	11·5
Fluency Scores:															
1. Names	—	491	241	148	136	610	233	209	-017	135	-057	-062	-125	-055	-007
2. Uses		—	363	421	452	470	595	447	253	390	072	058	002	019	001
3. Similarities			—	314	429	148	160	607	180	199	028	055	014	019	017
4. Pattern Meanings				—	741	-014	088	283	611	446	084	067	086	017	059
5. Line Meanings					—	107	241	266	381	632	119	090	084	087	114
Uniqueness Scores:															
6. Names						—	533	405	080	278	015	-096	-011	064	090
7. Uses							—	444	307	457	219	109	117	152	141
8. Similarities								—	212	315	027	047	058	097	111
9. Pattern Meanings									—	469	204	148	082	094	112
10. Line Meanings										—	192	191	044	025	002
Intelligence Scores:															
11. AL (verbal)											—	640	575	562	482
12. AQ (numerical)												—	305	388	319
13. Vocabulary													—	736	636
14. Speed of Comprehension														—	867
15. Level of Comprehension															—

Decimal points omitted. With d.f. = 122, critical values of the correlation coefficient are ·227 ($p < \cdot 01$), and ·175 ($p < \cdot 05$).

can be accounted for in terms of Pattern Meanings (fluency), or 29 per cent. of Uses (uniqueness) in terms of Names (uniqueness), the largest proportion of the variance of any creativity test which can be accounted for in terms of an intelligence measure is the 4 per cent of the variance of Patterns (uniqueness) which can be accounted for by AL.

However, the degree of internal consistency manifested by the creativity measures is very much less than that among the intelligence tests. This is particularly true of the uniqueness scores, where the median within-battery correlation of only ·361 is much smaller than the median intelligence test within-battery correlation of ·568. This may well reflect the reduced reliability of the creativity measures by comparison with the intelligence tests, as Wallach and Kogan (1965, pp. 41-44) report reliabilities of about ·8 for their tests.

The question of whether or not the two batteries of tests define separate dimensions of intellect was further tested by factor analyzing the combined matrix of intercorrelations. When principal axis factor analysis with unities in the diagonals was carried out on the data reported here, four factors with latent roots greater than unity were obtained (see Table 2). A large general factor accounted for 41·2 per cent of their common variance. This factor had high loadings on all fifteen variables, the smallest loading being ·315. A large second factor was also obtained (29·9 per cent of the common variance) for which loadings fell into two groups.

TABLE 2

PRINCIPAL AXIS FACTOR LOADINGS.

Variables	Factor				
	I	II	III	IV	h^2
Fluency Scores :					
1. Names	369	—366	—520	—079	547
2. Uses	731	—362	—211	—062	713
3. Similarities	588	—289	104	—594	793
4. Pattern Meanings	646	—202	587	—193	841
5. Line Meanings	690	—174	446	—042	707
Uniqueness Scores :					
6. Names	472	—272	—685	164	792
7. Uses	636	—100	—385	401	723
8. Similarities	617	—227	—200	—295	559
9. Pattern Meanings	557	—031	439	299	594
10. Line Meanings	667	—171	237	457	739
Intelligence Scores :					
11. AL (verbal)	413	690	007	212	692
12. AQ (numerical)	315	538	108	231	454
13. Vocabulary	332	755	—060	—161	709
14. Speed of Comprehension	376	810	—176	—184	862
15. Level of Comprehension	369	744	—176	—220	768
Sums of squares	4·32	3·13	1·85	1·18	10·48
% total variance	28·80	20·87	12·33	7·87	69·87
% common variance	41·22	29·87	17·65	11·26	100·00

Decimal points omitted.

In the case of the ten creativity variables, loadings ranged from —·031 to —·366 and were all negative, with a median of —·215. For the five intelligence tests, the range was ·538 to ·810 (median= ·744), and all loadings were positive. Thus, the factor analysis indicated the presence of a large general factor and a second, substantial bipolar factor of creativity *versus* intelligence.

The present findings may be summarized by saying that the Wallach-Kogan battery did indeed retain the property of internal consistency when administered to university students, while cross-correlations with the intelligence tests were usefully low. These results were obtained despite the fact that the tests were administered in a group form and so violated some of the authors' specifications for their use. However, it is by no means clear that they elicited a new and separate intellective mode. Although a substantial degree of separation of the two batteries was effected by factor analysis, a large general factor on which all tests loaded substantially was still obtained. Thus, provided that the presence of a considerable general factor is kept in mind, the Wallach-Kogan tests can usefully be used in the sampling of creativity, even with age levels other than that of grade five children, and in a group rather than individual form.

ACKNOWLEDGEMENTS.—The author is grateful to the Student Research Unit of the University of New England, Armidale, Australia, for their assistance in the conduct of this study.

IV.—REFERENCE.

WALLACH, M.A., and KOGAN, N. *Modes of thinking in young children.* New York: Holt, Rinehart and Winston, 1965.

(*Manuscript received, 30th October,* 1967)

THE VALIDITY AND RELIABILITY OF BENTLEY'S MEASURES OF MUSICAL ABILITIES

BY JOHN McLEISH
(*Cambridge Institute of Education*)

I.—SUMMARY.

A group consisting of twenty-five ESN children, aged from 8 to 15 years, fifty primary school children, aged from 9 to 10 years, and thirty secondary modern children, aged from 12 to 13 years, were tested with Seashore's, Bentley's and Wing's tests of musical ability.* The group was re-tested with Bentley's test exactly a year later. The children's scores on all the tests in the three batteries (six in Seashore, seven in Wing, four in Bentley) were positively correlated. A general factor was extracted, accounting for 35 per cent of the total test variance. This was identified as Burt's 'group factor of musical ability.' Bentley's measures show saturation coefficients of 0·646 on the average, in contrast to an average of 0·470 in Wing and 0·621 in Seashore. The test-retest reliability of Bentley's test (total score) was found to be 0·830. Of the three batteries, Bentley's was found to be easiest to use, causing least test anxiety in the three groups of children and being most economical in time needed for testing. The ESN group was found to be markedly inferior on all three batteries, less so on Bentley's tests than on the two others. However, in tests involving simple auditory discrimination (pitch, intensity) their average scores were equal to the other groups. But in more complex judgments, involving memory, their average scores dropped significantly. As far as the primary school group is concerned, the strong impression was given that not enough 'head-room' is provided, as the musical quotients (Bentley 'grades') were markedly bunched at the top end.

* The actual testing was carried out by Geoffrey Higgs, who knew most of the children in a teaching capacity. The study as a whole is reported in the reference below.

II.—REFERENCES.

BENTLEY, A. (1966). *Musical Ability in Children and its Measurement.* London: Harrap.

BENTLEY, A. (1966). *Measures of Musical Abilities.* London: Harrap.

McLEISH, J., and HIGGS, G. (1967). *An Inquiry into the Musical Capacities of E.S.N. Children.* Cambridge Institute of Education, Occasional Research Papers, No. 1, 12 pp.

(*Manuscript received 16th November,* 1967)

THE GROWTH OF THE SCHEMA OF PROPORTIONALITY

BY PETER PUMFREY

(*Manchester Child Guidance Clinic*)

SUMMARY. The development of the understanding of proportion by children of average ability aged from 5 to 15 years, was investigated using a modification of Piaget's clinical method. This facilitated an analysis of the strategies employed by the subjects in three problem situations involving proportional relations. The results suggest that an introduction to the understanding of proportion can profitably be made before the secondary stage of education, provided that the content and structure of the materials used are related to the strategies available to the child.

I.—AIM, METHOD AND SUBJECTS.

The main purpose of the investigation was to study the growth of the schema of proportionality in children between the ages of 5 and 15 years in the light of Piaget's theory of cognitive development. Three tests were adapted to be administered individually. Each test consisted of a series of situations in which the child is presented with three of the elements in a proportion and required to predict the fourth. Each of the tests used a different form of apparatus.

Test 1 (*Balance*). This test is based on the experiment of Inhelder and Piaget (1958, chapter XI). The apparatus consists of a simple balance and a set of weights. Following the procedure suggested by Jackson (1963), these took the form of sections of chain consisting of 1–9 links. The experimenter suspends a weight from a specified point on one arm of the balance and the subject is required to discover the point from which to suspend a similar or a different weight in such a way as to restore the equilibrium of the apparatus.

Test 2 (*Structural arithmetic material*). This test was based on preliminary investigations by Lunzer (Lunzer and Pumfrey, 1966). The apparatus consists of the standard set of Cuisenaire rods. The experimenter constructs a ' wall ' made up of a chain of rods of uniform colour and length. The rods are referred to as ' bricks.' The subject is required to predict the number of bricks of a different colour and length that he would need to construct the same overall length.

Test 3 (*Pantograph*). The child is required to predict the direction and amplitude of the movement of one pointer with respect to the other.

The computational aspect of the situation was kept as elementary as the material and proportions permitted. Standard forms of testing were used for all children. For each of the three tests a series of questions was prepared, similar in form, and involving a pre-arranged series of quantities and proportions. The children's successive predictions were recorded on prepared forms, and notes were also taken of the method by which these were reached. The procedure was aimed at combining the advantages of Piaget's clinical method with the requirement of objectivity in data collection (Flavell, 1963).

Four boys and four girls in each year from 5 to 15—a total of eighty children—were selected at random from a number of schools to which the writer had access in his capacity as an area remedial teacher to the Manchester LEA, with the restriction that no selected subject deviated by more than one standard deviation from the mean of each of two standardised tests.

II.—RESULTS.

The results of the three tests were examined on the basis of two criteria : (*a*) the accuracy of the children's predictions, regardless of the method by which these were achieved ; (*b*) the method of reasoning, or the ' strategy ' used in each prediction.

The reliability coefficients (split-half) of the three tests in terms of the first criterion were : Test 1, ·960 ; Test 2, ·928 ; Test 3, ·953. The relation of chronological age to accuracy of prediction was : Test 1, ·562 ; Test 2, ·684 ; Test 3, ·898. The relation of mental age to accuracy of prediction was : Test 1, ·615 ; Test 2, ·739 ; Test 3, ·918. Inter-correlations for accuracy of prediction for the three tests were : Tests 1 and 2, ·401 ; Tests 1 and 3, ·599 ; Tests 2 and 3, ·627. All these correlation coefficients are significant at the ·01 level. None of the differences between correlations for chronological and mental age for a given test is significant.

Partial correlation was used to hold constant the effects of the relationships between chronological age and accuracy in prediction and between mental age and accuracy. With chronological age partialled out, only the correlation between Test 1 (Balance) and Test 3 (Pantograph) remained significant at the ·05 level. With mental age partialled out, none of the inter-correlations remained significant. Thus, the effect of partialling out the larger correlation was in the anticipated direction, and the construct validity of the tests was not supported by the resultant inter-test correlations to the extent anticipated. This lack of support may be the result of the arbitrary definition of ' accuracy in prediction ' adopted as a basis for the calculations, the homogeneous characteristics of the sample and the specificity of the tests.

Test 2 (Structural arithmetic) was most readily dealt with by the children. Test 3 (Pantograph) was intermediate in difficulty ; and Test 1 (Balance) was found most difficult by the majority of the children tested. This order of difficulty was found to hold not only for the sample as a whole but also for most of the separate age groups from 5 to 15.

While the growth in the accuracy of prediction in each of the three test situations provides an unambiguous background of reference for evaluating the development of proportionality, the validity of Piaget's interpretation cannot be assessed without a more searching analysis, even if this entails some reduction in objectivity. A number of writers have criticised Piaget's own method of analysis, which is an overall interpretation based on an assessment of each subject's protocol taken as a whole (Flavell, 1963, pp. 430-33). Accordingly, it was decided to obtain unambiguous 0 or 1 assessment on each of a relatively large number of categories, as applied to each separate prediction in the three tests. Such categories could then be taken either singly or in combination to serve as criteria for the determination of strategies.

After examination of the protocols, approximately twenty-five categories were defined, the number varying slightly for each of the three tests. By way of example, the following categories were used among the twenty-six applied to the data for Test 1 (Balance) :

Interferes with balance : the child refuses to accept the objectivity of the situation.

Differences : the child appreciates need for differences in position of different weights to maintain equilibrium.

Compensates : the child appreciates that direction of difference in distance of W2 from the fulcrum is such as to compensate for the direction of the difference in W1 and W2.

Simple addition : addition to one aspect of the situation only (e.g., to length without reference to weight).

Proportion by symmetry : appropriate to special cases, e.g. : $\underset{L}{3} \times \underset{W}{4} = \underset{W}{4} \times \underset{L}{3}$

Set continued though inadequate : the child persists in applying a technique inappropriate to the situation, e.g., doubled distance because doubling was successful immediately before.

Reverts to old strategy : the child changes strategy and uses, inappropriately, one which had been used in previous items.

Attempts new strategy : the child attempts a new strategy, which may or may not be correct, but the prediction includes actions or explanations not offered before.

Operational proportionality : the child multiplies distance by weight on one arm and divides by weight on the other to obtain desired distance, and can generalise rule.

The reliability of the assessment was tested by having a ⅛ sample of the data from each test categorised separately by two markers. A satisfactory overall percentage agreement of between 96 per cent and 97 per cent was found for each of the three tests, though in terms of the categories within each test the percentage agreement varied between 50 per cent and 100 per cent.

The percentage agreement for the category 'Attempts new strategy' was 95 per cent. for Test 1 (Balance), 90 per cent for Test 2 (Structural arithmetic) and 85 per cent for Test 3 (Pantograph). The means and standard deviations of the incidence of new strategies used in the three tests were, respectively, 4·5 and 2·2 for Test 1; 4·1 and 1·1 for Test 2; and 5·1 and 2·5 for Test 3. Comparison of the growth curves of the incidence of 'Attempts new strategy' with age group, shows that the perceptually simplest test, Test 2, produced the lowest maximum of new strategies. Test 3 gave a higher maximum of new strategies; and Test 1 the most complex of the tests, produced the highest maximum of new strategies. The children's ability to produce new strategies in Test 1 reached a maximum at the age of 12–13 years, a finding in agreement with Jackson (1963, p. 219).

A previous enquiry (Lunzer and Pumfrey, 1966) had established three clearly defined strategies in the prediction of proportions using structural arithmetic apparatus. These were, respectively, the numerical equivalence of elements and the addition or subtraction of differences, the iteration of the lowest common multiple match and the use of a compound equivalent unit. All these strategies were found in Test 2 and the first two were identified in the responses to Test 3. These strategies were found to be related to the perceptual complexity of the problem situation, the particular type of proportion being investigated, the age of the child and the nature of the material. In all three tests additive strategies were gradually replaced by additive-multiplicative strategies. The adequacy of the strategy in general increased with age.

The percentage of each age group exhibiting operational proportionality in each test, was found to vary. Test 1 was the first test in which its application was manifest, followed by Test 3. Test 2 was not dealt with at the formal level.

III.—Conclusions.

The fall off in the incidence of new strategies from the age of 12–13 years coincided with an increase in Test 2 and Test 3 of the children's predictive ability of more difficult ratios than 2 : 1, 3 : 1 and their inverses, using multiplicative strategies. Simultaneously, their quantitative appreciation of the function of the balance showed a marked development. Thus, the decline in new strategies occurred at the age when the children were beginning to make consistently correct predictions using the schema of proportionality. Therefore, while appearing less flexible in terms of the number of new strategies used, the children's thinking was starting to function in a new mode, far more effective than its previous one.

The essential difference in difficulty to the child between situations involving both proportion and reciprocity in a physical system and those involving only proportion was supported by the data.

The findings suggest that an introduction to the understanding of proportion can be made before the secondary stage of education, provided that the structure and content of the materials used are related to the limitations in strategy, and hence in understanding, available to the child. In this respect, the findings are in agreement with Winch (1913).

Acknowledgement.—The author is indebted to Dr. E. A. Lunzer for his constructive criticism during all the stages involved in the preparation of this note.

IV.—References.

Flavell, J. H. (1963). *The Developmental Psychology of Jean Piaget.* New York: Van Nostrand.

Inhelder, B., and Piaget, J. (1958). *The Growth of Logical Thinking.* London: Routledge.

Jackson, S. (1963). *The Growth of Logical Thinking in Normal and Subnormal Children.* Unpublished M.Ed. thesis, Univ. of Manchester.

Lunzer, E. A., and Pumfrey, P. D. (1966). Understanding proportionality. *Maths. Teach.*, 34, 7-12.

Winch, W. H. (1913). Should young children be taught Arithmetical Proportion? *J. of Exp. Ped.*, 2, 79-88, 319-329, 406-420.

(*Manuscript first submitted 29th November, 1966*)

VALIDATION OF A SHORT FORM WISC WITH CLINIC CHILDREN

By D. W. McKERRACHER and R. A. WATSON

(*Rampton Hospital, Retford, Nottinghamshire*)

Summary. One-hundred-and-twenty-seven children tested at a child guidance clinic were given a full scale WISC in the course of their examination. Prorated IQs were calculated from four of the ten sub-tests (Arithmetic, Vocabulary, Picture Arrangement and Block Design). These prorated IQs correlated +0·96 with full scale IQ and gave accurate indications of level of performance on the full scale. This shortened version is least suited for young children and children in the IQ range 70-89.

I.—Introduction.

Since the publication of the Wechsler Adult Intelligence Scale (1955), sporadic attempts have been made to investigate the validity of Doppelt's claim (1956) that the full scale IQ could be predicted in a normal population with reliability and validity from the administration of only four subtests (Arithmetic, Vocabulary, Block Design and Picture Arrangement). Clayton and Payne (1959) found correlations ranging from ·92 to ·97 in a clinical population composed of various types of mentally disturbed adults, including organic, functional and personality disorders. McKerracher, *et al.* (1967) demonstrated similar correlations in the subnormal and psychopathic inmates of a special security hospital.

Previous attempts to find useful abbreviated versions of the WISC have been mainly confined to samples of children in institutions for mental retardation (Carleton and Stacey, 1953) or in special classes for the retarded (Finley and Thompson, 1958). Both studies obtained Pearson correlations in excess of ·8, but *five* subtests had to be employed. In the only clinical population similar to the subjects investigated in the present study, Yalowitz and Armstrong (1955) found Pearson correlations of only ·5 between their short-form estimates and the WISC full-scale IQs of children referred for various reasons to a child guidance clinic. They concluded that their abbreviated version was invalid.

All of these correlations are lower than those found in adult populations using Doppelt's method of estimation from the WAIS. Without resorting to a regression analysis of the correlations between each individual WISC subtest and the full-scale score, it is possible on an empirical basis to approximate the full score by simple proration, using the same subtests as selected by Doppelt from the adult test: (Arithmetic + Vocabulary + Picture Arrangement + Block Design) × 10/4 = estimated Full Scale Score.

This measure is obviously not so satisfactory as selecting the subtests which are most highly correlated with the full-scale score from the WISC data by a full regression analysis. Without computer facilities, such an untertaking was not possible in the time available for this project. It might be expected with children that the inclusion of Arithmetic would reduce the predictive accuracy of a short-form version, since it is more likely to be sensitive to the effects of disturbed schooling. However, the findings discussed below do not appear to support this hypothesis as there was no indication of an underestimate of the full-scale IQ. It is also difficult to believe that any other combination of subtests would markedly improve the obtained degree of correlation with the full scale IQ.

II.—Method.

The psychological files in a Lindsey Child Guidance Unit were examined, and in every case where a full-scale WISC (English Version, 1949) had been administered, a prorated IQ was calculated for comparison with the actual IQ. The average age of this sample (N = 127: 85 boys and 42 girls) was 8·56 years (S.D. = 2·26). The average WISC Full Scale intelligence was 89·07 (S.D. = 17·4).

Reasons for referral to the clinic were varied: Brain-damaged, psychotic or genetically abnormal, 6; enuresis or encopresis, 3; anti-social behaviour problems (non-criminal), 13; anti-social behaviour problems (criminal), 9; emotionally maladjusted at home and schoool 13; emotionally maladjusted with retardation, 8; learning difficulties due to physical abnormalities (deafness, stammering, speech defects, etc.), 7; in need of ESN schooling, 33; anomalous learning difficulties without emotional maladjustment or backwardness, 35.

III.—RESULTS.

Table 1 shows a highly significant correlation of + ·96 between the prorated and the full scale IQ. It is interesting that this degree of relationship with the Full Scale IQ is greater than those with the Verbal and Performance halves considered separately. This suggests that the prorated IQ is an effective means of arranging the intelligence scores of emotionally disturbed children in much the same order of merit as that of the full test. It does not automatically follow that the prorated IQ provides a reasonable estimate of the full IQ score value in absolute terms. Close agreement about order of precedence could still exist though the means and standard deviations of each technique were grossly dissimilar. However, it can be seen from the data in Table 1 that there is very little discrepancy between the IQ estimates in terms of average score values and deviations around the means. The standard error of estimate was computed and amounted to 5·69 IQ points. The prorated IQ should, therefore, provide an acceptable indication of the functional intelligence of a subject in score level, as well as in order of ability.

TABLE 1

MEANS, STANDARD DEVIATIONS AND CORRELATIONS: WISC FULL SCALE AND PRORATED IQs.

	Verbal	Performance	Full Scale	Prorated
Mean	89·39	90·62	89·07	89·55
S.D.	17·90	20·22	17·41	20·03

No significant differences.

Correlations	Verbal	Performance	Prorated
Full Scale	·935	·937	·960
Verbal	—	·754	·889
Performance	—	—	·658

All correlations significant (p< ·01).

To check on this assumption, the actual differences between the obtained and estimated full-scale IQs were arranged (*a*) in age groups; and (*b*) in classifications of intellectual level according to the full scale WISC IQ.

The method was seen to be least effective in the 5 to 7-year-old group where only 59 per cent of the prorated IQs were within ±5 points of the full scale IQs (see Table 2). But between 67 and 75 per cent of the children aged 8 to 13 years, and *all* of the 14-year-olds or over, fell within this area of estimation.

Taking them to be the most reliable and valid estimates of ability, the full-scale IQs were arranged in the categories recommended by Wechsler (1958) for classification purposes. The full IQs were regarded as the constants, and the differences between them and the prorated IQs calculated (see Table 3). Only one error of selection was noted among children scoring in the subnormal range (full-scale IQ 69 and below). A higher percentage of discrepancy was present in the full-scale IQ range 70 and over, using ±5 points as the criterion of concordance. The least effective prorated estimates were for the full-scale IQs in the borderline defective to dull average range (70–89). Only 62 per cent. were within ±5 points of the actual IQ.

TABLE 2

DIFFERENCES BETWEEN PRORATED AND FULL SCALE IQS. IN AGE RANGES.

Differences	Years			
	5—7	8—10	11—13	14+
+16—20	1 (3)	0 (0)	0 (0)	0 (0)
+11—15	2 (6)	2 (4)	2 (6)	0 (0)
+6—10	7 (22)	6 (13)	2 (6)	0 (0)
+1—5	10 (31)	15 (31)	12 (33)	1 (9)
0	1 (3)	4 (8)	1 (3)	3 (27)
—1—5	8 (25)	18 (36)	11 (31)	7 (64)
—6—10	3 (10)	4 (8)	6 (18)	0 (0)
—11—15	0 (0)	0 (0)	0 (0)	0 (0)
—16—20	0 (0)	0 (0)	1 (3)	0 (0)
Totals	32 (100)	49 (100)	35 (100)	11 (100)

The Full Scale IQs are regarded as the constants and the differences between them and the prorated IQs are shown as plus or minus to indicate whether the short-form provides an over- or under-estimate. (Percentages in brackets.)

TABLE 3

DIFFERENCES BETWEEN PRORATED AND FULL SCALE IQS IN INTELLECTUAL CLASSIFICATION RANGES BASED ON FULL SCALE IQS.

Differences	Full Scale IQs.			
	69—	70—89	99—109	110+
+16—20	1 (5)	0 (0)	0 (0)	0 (0)
+11—15	0 (0)	2 (4)	3 (7)	1 (5)
+6—10	0 (0)	7 (15)	5 (13)	3 (15)
+1—5	7 (39)	12 (25)	12 (30)	7 (33)
0	3 (17)	3 (6)	3 (7)	0 (0)
—1—5	7 (39)	15 (31)	14 (35)	8 (37)
—6—10	0 (0)	9 (19)	2 (5)	2 (10)
—11—15	0 (0)	0 (0)	0 (0)	0 (0)
—16—20	0 (0)	0 (0)	1 (3)	0 (0)
	18 (100)	48 (100)	40 (100)	21 (100)

(Percentages in brackets)

IV.—DISCUSSION.

Pressure of work in a child guidance clinic often makes the administration of the full-scale WISC impossible. In cases where intelligence is not considered to be of primary importance, a reliable and valid indication of ability using a less time-consuming technique is much to be desired, as an alternative to the 'starred' Binet method (Terman and Merrill, 1961). The prorated IQ based on four WISC subtests seems to meet this requirement with some success. It is capable of selecting mentally subnormal children with a minimal amount of error, and is adequate at estimating the ability of subjects of normal, or above normal intelligence. It is weakest at estimating the IQs of young children (below 7 years); and children of any age in the intellectually very dull category (IQ 70-89).

Two further reservations must be made: (*a*) the prorated IQs were calculated from four subtests, extracted after all ten WISC subtests had been administered (Digit Span and Mazes excluded). It is possible that the application of the abbreviated form on its own outside the context of the full WISC, might produce slightly different results. Only a follow-up and retest with a full WISC could provide an answer to this criticism; (*b*) even with a correlation as high as ·96, errors of selection are to be expected. Shapiro (1951) explained that with a test-retest correlation of ·9 and a standard deviation of 16, out of every ten children obtaining an average score of 100 on the first test, one child would obtain a score above 112 or below 88 on the second test. This comment applies equally well to the relationship between the prorated and full scale IQs.

V.—References.

Carleton, F. O., and Stacey. C. L. (1953). Evaluation of selected short forms of the Wechsler Intelligence Scale for children. *J. Clin. Psychol.*, 9-10, 258-261.

Clayton, H., and Payne, D. (1959). Validation of Doppelt's WAIS short form with a clinical population. *J. Cons. Psychol.*, 23, 467.

Doppelt, J. E. (1956). Estimating the full scale score on the WAIS from scores on four subtests. *J. Cons. Psychol.*, 20, 63.

Finley, C. J., and Thompson, J. (1958). An abbreviated Wechsler Intelligence Scale for children for use with educable mentally retarded. *Am. J. Ment. Defic.*, 63, 473-480.

McKerracher, D. W., Watson, R. A., Little, A. J., and Winter, K. S. (1967). Validation of a short form estimation of WAIS in subnormal and psychopathic patients. *J. Ment. Sub.* (in press).

Shapiro, M. B. (1951). An experimental approach to diagnostic testing. *J. Ment. Sci.*, 97, 748.

Terman, L. M., and Merrill, M. A. (1961). *Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M.* Boston: Houghton Mifflin.

Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children.* New York: Psychol. Corp.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale.* New York: Psychol. Corp.

Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence.* Baltimore: Williams and Wilkins.

Yalowitz, J. M., and Armstrong, R. G. (1955). Validity of short forms of the Wechsler Intelligence Scale for Children (WISC). *J. Clin. Psychol.*, 11, 275-277.

(*Matnuscri preceived* 12*th January*, 1968)

A FIVE-YEAR FOLLOW-UP OF SECONDARY SCHOOL SELECTION PROCEDURES IN CENTRAL AFRICA, 1962-1967

By S. H. IRVINE

(*Institute of Education, University of Bristol and Educational Testing Service, Princeton*)

I.—Background.

This study was begun in 1962 to provide more information about the nature and function of the selection system operating in Rhodesia. At first, only limited follow-up was considered possible, but the study has been extended to follow the progress of African pupils systematically through the secondary school system. A parallel study was begun in Uganda at the same time by Mr. J. S. Silvey, of the University of Nottingham, and the possibility of comparison of two systems of education has been realised through close co-operation and detailed planning. This paper concentrates on central African work with two main aims: (*a*) the provision of information about the kind of tests used and their utility; (*b*) providing an outline of methods used so that future studies may evaluate the model implied.

It is, perhaps, appropriate to indicate additional sources of information about these studies at the outset so that they can be followed in detail if necessary. The central African studies may be found in Irvine (1964, a, b, c, ; 1965 a, b,) and the East African studies in Silvey (1963, 1964). Other relevant work in selection in developing countries may be seen in Vernon (1961) and Wantman (1962), Schwarz (1961, 1963), and Ferron (1965). Two important bibliographies are Hopkins (1962) and Andor (1966).

II.—THE FIRST SURVEY.

A survey of educational skills and abilities was undertaken in 1962, towards the end of the eighth year of education, just prior to the secondary school selection examinations. All students took a battery of tests and completed a personal record card. The tests, with all others used at various levels, are included in Appendix A. The experimental tests were ability tests in English, a test of reasoning in figural 'non-verbal' context, comprehension, vocabulary and spelling tests. All required special testing procedures. These were combined with the results of the traditional achievement examination and analysed. The sample consisted of 1,842 students from the schools of a province in Rhodesia selected because its examination results fitted closely national norms. One-third of all schools, chosen at random from stratifications by type, quality and geographical location, were verified as fitting provincial and national patterns of selection examination results in 1962.

The analysis of this survey included the provision of base-line norms for the experimental tests, estimates of concurrent validity with scaled head masters' estimates as the criterion, and factor analysis of test content. From the results, the following conclusions were drawn :

(1) Experimental tests could add significantly to the concurrent validities (R= ·89) across all types of schools.

(2) The main dimensions of ability were reasoning, verbal skills and number skills ; and the traditional selection system included tests redundant for prediction purposes since too many measured the same components.

(3) The major influence on the scores was the quality of school attended. Also, males did significantly better than females on verbal tests, especially those demanding information about scientific or social studies. Age (as reported, but not verified since births were not registered by law) was negatively correlated with attainment, even after holding number of years of education constant. Analyses of variance showed little or no significant association of scores with level of father's occupation, number of siblings, family position, or number of languages spoken.

These results were largely replicated the following year in Zambia with the same basic design and experimental tests (see Irvine, 1964b).

Stage One Follow-Up, 1963-1964. The utility of the tests was further investigated in 1963 by constructing an achievement battery of objectively-scored tests in consultation with principals of secondary schools and officials of the Ministry of Education. They were on 'core' subjects of English language and comprehension, arithmetic processes, algebra, geometry and general knowledge. Headmasters' estimates of performance were scaled on the total of these tests to broaden the base of criterion. All the students (just over 1,000) in the first year of the provincial secondary schools were tested, and of these 291 were available among those identified as belonging to the original sample. Tests of goodness of fit showed the follow-up sample to be representative of the population from which it was drawn. Analysis showed that the multiple correlations between the traditional examination and the final global criterion of Form 1 total plus scaled headmasters' estimate could be significantly increased by inclusion of ability tests and primary school headmasters' estimate. Once again, the quality of secondary school exerted a strong influence on performance. Pupils rated equal in the selection examinations tended to be differentiated if they went to schools of high, middle, or low quality.

TABLE 1

PREDICTING ACHIEVEMENT IN 1964 JUNIOR CERTIFICATE (N=231).

J.C. Subjects	1962 Tests		1963 Tests		All 1962, 1963	
	No. Used	R	No. Used	R	No. Used	R
English	3	·46	3	·59*	3	·60
Arithmetic	4	·59	3	·62	5	·67
Mathematics	3	·48	5	·74*	5	·75
History	1	·32	4	·36	4	·42
Geography	3	·41	2	·35	6	·48
Vernacular Language	4	·50	2	·56	3	·58
Pass/Fail	2	·46	4	·58*	7	·63

* Significantly different from 1962 R. Under the " number of tests used " column is shown the number of tests required to produce the multiple correlation from the total battery available (14 tests in 1962, 6 in 1963, 20 tests in all).

It was possible in this phase of the investigation to trace the sample through to the first operational criterion in the survey, the Junior Certificate national examination administered at the end of two years of secondary education. Numbers were by this time reduced to 231, a wastage of nearly one-fifth of the total identified in Form 1. The summary of results is included in Table 1. Each test used in 1962, traditional or experimental, was used to make up a battery for multiple regression on core subjects in the examination and on a pass/fail dichotomy. The norms of all the 1962 tests were revised for the 231 pupils involved. The 1963 Form 1 tests were treated similarly and the two sets combined to produce multiple correlations for all available test data. The results showed that the number of tests required to produce a multiple correlation significant at the ·01 level of probability was considerably smaller than the number of tests actually used. Second, the most consistent predictors of success were ability tests and scaled headmasters' estimates. Third, there were significant changes in subject content from primary to secondary schools that were reflected in the correlations for two key subjects, English and Mathematics. However, factor analysis again showed that the major constructs in the tests were reasoning, verbal and numerical skills. Male aptitude was still present in informational subjects but not nearly so pronounced as in the primary school.

Stage Two Follow-Up, 1966. In 1966, the pupils took the final hurdle in the secondary system, the secondary school leaving certificate examination. Just prior to this, in September, the writer revisited the schools and tested every pupil in them, a total of 342. Of these only 80 had survived from the original 1,800 tested in the primary schools in 1962. A battery of tests of vocational aptitude was administered, a detailed personal questionnaire on background, vocational interests and aspirations was also given. Headmasters' estimates of success in the forthcoming examination were collated and a follow-up questionnaire on all student drop-outs completed. The data are being analysed at the time of writing and will be related to examination success. Secondly, a detailed study was completed in one secondary school using 30 of the French (1963) battery of factorial tests to give insight into the interaction of such tests across cultures.

Stage Three Follow-up, 1967. A postal follow-up of all 342 pupils was begun in April 1967 to find out what had happened to them on leaving school. By October, 272 replies to the inventory sent out had come back, representing 80 per cent return. For postal follow-up, this was regarded as a high figure and the relative success of this part of the survey was held to be due in part to the effort made to give individual test results to the students concerned. First analysis shows that of 272 respondents, 113 are in further full-time education, 39 are employed and 120 are unemployed.

The analyses proposed are as follows:

(1) Regression on school certificate results using the groups of tests available in 1962, 1963, 1964 (Junior Certificate) and the experimental survey tests of 1966.

(2) The relation of cognitive variables to occupational aspirations.

(3) A search for noncognitive predictors of intellectual capacity.

(4) Study of the dimensions of occupational prestige and work preferences.

(5) Study of job aspirations in relation to realistic expectations.

III.—Conclusion.

What began as an exercise on selection efficiency has now turned into a long-term study of the development and use of high-level African manpower. The opportunities are unique and the information likely to prove fruitful for planners and administrators concerned with underdeveloped countries. When the research in Uganda has been joined with the work in central Africa, the comparisons are likely to be sociologically as well as educationally illuminating.

Acknowledgement is made of financial help from the University of Rhodesia and Nyasaland Research Fund, The Ministry of Education in Rhodesia, The Anglo-American and Rhodesia Selection Trust groups of mining companies, the Ministry of Overseas Development and the Leverhulme Trust, whose senior research award enabled the writer to return to Africa in 1966 for field work. This paper was first presented at the International Testing Conference in Berlin, May, 1967, and was revised during tenure as Visiting Scholar in the Personality and Social Research Group at ETS.

IV.—References.

Andor, E. (1966). *Abilities of the African in Sub-Saharan Africa*, 1784-1963. N.I.P.R., Johannesburg.

Ferron, O. (1965). The test performance of 'coloured' children. *Educ. Res.*, 8, 42-57.

French, J. W. (1963). *Kit of Reference Tests for Cognitive Factors* (*Rev.*). Princeton, N.J.: Educational Testing Service.

Hopkins, J. (1962). Bibliographie des recherches psychologiques conduites en Afrique. *Revue de Psych. Appl.*, 12, 201-213.

Irvine, S. H. (1964a). *A Psychological Study of Selection Problems at the end of Primary Schooling in Southern Rhodesia*. Unpublished Ph.D. thesis, Univ. of London Library.

Irvine, S. H. (1964b). *The Northern Rhodesia Mental Ability Survey* (with R. S. MacArthur and A. R. Brimble). Zambia: Rhodes Livingstone Institute.

Irvine, S. H. (1964c). Selection of Africans for post-primary education in Southern Rhodesia. Pilot survey, June-July, 1962. *Bull. Inter-African Lab. Inst.*, 11, 69-93.

Irvine, S. H. (1965a). Selection for secondary education in Southern Rhodesia. *Faculty of Education, Paper No. 4.* Salisbury, Rhodesia: University College.

Irvine, S. H. (1965b). Adapting tests to the cultural setting: a comment. *Occup. Psychol.*, 39, 12-23.

Schwarz, P. (1961). *Aptitude Tests for Use in the Developing Nations.* Pittsburg: Amer. Inst. Res.

Schwarz, P. (1963). Adapting tests to the cultural setting. *Educ. Psychol. Meas.*, 23, 673-686.

Silvey, J. (1963). Aptitude testing and educational selection in Africa. *Rhodes Livingstone J.*, 34.

Silvey, J. (1964). Selection for senior secondary schools in Uganda. E.A.I.S.R. Kampala, Uganda (mimeographed).

Vernon, P. E. (1961). *Selection for Secondary Education in Jamaica.* Govt. Printer, Kingston.

Wantman, M. J. (1962). *Report on the Malayan Project in Educational and Occupational Selection in West Africa* (Ed. Taylor). London: O.U.P.

(*Manuscript received 8th February, 1968*)

V.—APPENDIX A.

TESTS AND INVENTORIES USED 1962-1966—CENTRAL AFRICA.

NIPR—National Institute for Personnel Research, Johannesburg.
NB—Normal Battery.
Std. 6—Traditional Selection Examination.
AIR—American Inst. for Research ID Tests, Pittsburgh, U.S.A.

1962 (8 Years of Education, N=1842)		
Tests		Estimates, Inventories
Raven 38.	Std. 6 English 1.	Headmasters.
NIPR Spiral Nines.	Std. 6 English 2.	School Quality.
NIPR NB Comprehension.	Std. 6 Arithmetic (Mech.).	Record Card.
NIPR NB Vocabulary.	Std. 6 Arithmetic (Problems).	
NIPR NB Spelling.	Std. 6 Arithmetic (Mental).	
	Std. 6 History/Cur. Affairs.	
	Std. 6 Geog./Nature Study.	
1963 (9 Years of Education, N=1030)		
Tests		Estimates, Inventories
Irvine English.	Irvine Algebra.	Headmasters.
Irvine Arithmetic.	Irvine Geometry.	School Quality.
Irvine General Knowledge.		
1966 (12 Years of Education, N=342)		
Tests		Estimates, Inventories.
AIR Verbal Analogies Higher.	AIR Figures.	Headmasters'
AIR Reading Higher.	AIR Names.	School Quality.
AIR Graphs.	AIR Codes.	Personal Data.
AIR Tables.	AIR Mech. Information.	
AIR World and Science Information.		

THE FEMININE IMAGE FACTOR IN GIRLS' ATTITUDES TO SCHOOL SUBJECTS

BY F. W. SLEE

(*Chorley College of Education*)

SUMMARY. Differences in the structure of attitudes to school subjects between secondary modern school boys and girls were revealed by a factorial analysis of subject preferences. While it was possible to describe the boys' attitudes in terms of subject content, the structure of the girls' attitudes was more complex. Of importance was a factor associated with the girls' perception of their future role. This was identified as being similar to the 'feminine image' factor previously described by Hallworth and Waite.

I.—INTRODUCTION.

In two papers Hallworth and Waite (1963 and 1966) have used the semantic differential technique to investigate the value judgments of 14- and 15-year-old boys and girls. The subjects' attitudes were measured by asking them to judge a number of concepts such as teachers, failure, money, etc., by rating them on a series of evaluative scales. After correlating the scores derived from these scales for all concepts and factorising the resulting matrix, Hallworth and Waite obtained a description of the evaluative space of the adolescents concerned. Of considerable prominence and interest among the factors extracted from the value judgments of the girls was one named by the authors 'the Feminine Image.' This well-defined factor they described as being "apparently related to ambitions and pleasures which are regarded as typically feminine and with which it would seem that these girls tend to identify themselves" (Hallworth and Waite, 1963, p. 43). As may be expected this concern with a clearly identifiable sex role was not so obvious among the boys who were interested in a more general group of concepts such as ambition, success and reward. Obviously, these differences in perception between adolescent boys and girls demonstrated by Hallworth and Waite have considerable educational significance. It is the object of this research note to offer further evidence regarding these differences and to show how the differing perception by boys and girls of their future role influences attitudes to school subjects.

II.—THE INVESTIGATION.

A new Thurstone-type generalised attitude scale of twenty-five items to measure attitude to school subjects was constructed and administered to 265 boys and 248 girls in the 12+, 13+ and 14+ age groups in seventeen classes in nine different schools. These 513 children were a sample drawn as representative of a population of children in maintained secondary modern schools in a Lancashire County Borough. In drawing the sample the class was used as the sampling unit and the sample was stratified by age, and by size of school.

Attitude scores were collected for eleven subjects namely, art, physical education, science, geography, history, English, religious education, mathematics and music, together with handicraft (woodwork or metalwork) and technical drawing for the boys and housecraft and needlework for the girls. In addition, practical and academic interest scores were obtained from all the children on the Devon Interest Test (Fitzpatrick and Wiseman, 1955), together with E and N scores on the Junior Maudsley Personality Inventory (Furneaux and Gibson, 1961). Scores on a scale measuring attitude to school (Fitt, 1956) and the age of the children completed the seventeen variables in the battery.

III.—STATISTICAL ANALYSIS.

The boys' and girls' scores derived from the testing were treated separately. As attitude scores tend to give skewed distributions, all sets of scores were first normalised, then inter-correlated. For each correlation matrix a Principal Components analysis was carried out and in both the analyses six significant factors were extracted and rotated to the Varimax criterion. After Varimax rotation, the percentage variances of these factors for the boys were 9·7, 10·4, 11·4, 9·2, 11·0, and 7·2, respectively (total variance 58·8 per cent). For the girls the percentage variances were 14·7, 12·8, 9·4, 8·7, 6·6 and 6·3, respectively (total variance 58·5 per cent).

A full Promax analysis was then carried out and significant second order factors were extracted, two in the case of the boys' and three in the girls' analysis. These second order factors were again rotated to the Varimax criterion. Tables showing the full details of the statistical analysis are available from the author.

IV.—RESULTS AND DISCUSSION.

(*a*) *The Boys' Analysis.*

The first order factors in the boys' analysis have been tentatively identified as follows:

I. A human interest factor showing a dislike for geography, history and, to a lesser extent, religious education, especially among older boys.

II. A scientific-mechanical factor with loadings on science, handicraft and technical drawing. This factor is associated with stable introversion. A loading on English is difficult to account for here.

III. A Devon Interest Test Factor with highest loadings on the test itself. Favourable attitudes to handicraft and technical drawing are associated with practical interest as shown by the test.

IV. A factor showing liking for subjects involving physical and manipulative activities such as art, physical education, handicraft and technical drawing.

V. A factor showing dislike for abstract or academic subjects such as mathematics, religious education and history.

VI. A personality factor associated with stability and older boys. The only subject loading on this factor shows an unfavourable attitude to music.

Factors I, II, IV and V may be interpreted as showing that the response of the boys to school subjects was on a subject content basis. This interpretation is supported by the second order factor pattern which splits the subjects into practical and academic groupings, the former containing handicraft, art, physical education and science and associated with stability and the latter appealing more to young boys

(*b*) *The Girls' Analysis.*

The girls' first order factors may be tentatively described as follows :

I. A human interest factor associated more with younger introverted girls. Loadings occur in religious education, history and music. The latter may be explained by the fact that in many schools the subject includes the history of music.

II. The feminine image factor showing a liking for housecraft, art, English and music, which is strong, particularly among older girls.

III. A Devon Interest Test Factor showing practical interest associated with liking for needlework and academic interest linked with liking for English and geography.

IV. A factor showing liking for subjects involving physical and manipulative activities, such as P.E. and needlework. A loading on geography here is difficult to explain.

V. A mathematical-scientific factor with high loadings on those two subjects only.

IV. A personality factor of stable introversion showing a liking for music as the only loading among the subject variables.

The parallel with the boys' analysis is obvious except that the girls' factor II shows a strong liking for housecraft and other subjects which are often regarded as appropriate to a woman's education. It is this factor which it is suggested may be equated with the feminine image factor described by Hallworth and Waite. The effect of this clear conception by the girls of the feminine role is also evident in the second order factors. Here the first factor is one of interest in academic subjects such as history, English, science, and religious education, which is associated with introversion ; the second shows a liking in particular for needlework and physical education associated in this case with stability ; and finally, the feminine image factor showing a liking for housecraft, art and music, particularly among the older girls. The boys' conception of their sex role is apparently far more generalised than in the case of the girls as a masculine image factor is not readily identifiable in the boys' analysis.

V.—CONCLUSIONS.

The usual classification of subjects according to content as academic or practical has some validity when used to describe the attitudes of secondary modern school boys to subjects in school. But this classification is not so useful when describing girls' attitudes to school subjects as their attitudes are strongly influenced by a 'feminine image' factor previously described by Hallworth and Waite and reflecting the girls' conception of their feminine role. The importance of this role factor in planning courses for adolescent girls is obvious. Whether the scientific-mechanical factor stands in the same relationship to boys' interests as the feminine image factor does to the girls' is a matter for further research.

VI.—REFERENCES.

FITT, A. B. (1956). An experimental study of children's attitudes to school in Auckland, N.Z. *Brit. J. Educ. Psychol.*, 26, 25-30.

FITZPATRICK, T. F., and WISEMAN, S. (1955). *The Devon Interest Test.* Edinburgh : Oliver and Boyd.

FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.

HALLWORTH, H. J., and WAITE, G. (1963). A factorial study of value judgments among adolescent girls. *Brit. J. Stat. Psychol.*, 16, 37-46.

HALLWORTH, H. J., and WAITE, G. (1966). A comparative study of value judgments of adolescents. *Brit. J. Educ. Psychol.*, 36, 202-209.

(*Manuscript received 17th January,* 1968)

BOOK REVIEWS

ALEXANDER, F. G., and SELESNICK, S. T. (1967). *The History of Psychiatry.* London: Allen and Unwin, pp. xvi+472, 63s.

Franz Alexander, who died in 1964, was well known as a leading American psychoanalyist. Born in Budapest in 1891, he was the first psychoanalytic trainee to undergo what has since become the regular training procedure. After settling in America he did much to organise and advance psychoanalysis in that country. The present book is to a considerable extent Alexander's work. The first draft was completed in collaboration with Dr. Selesnick, a student and later colleague of Alexander, before the latter's death. Not surprisingly, the book is basically psychoanalytic in its orientation, though it is inclusive enough to embrace most approaches to psychiatry in its purview.

The book begins in the ancient world with Mesopotamia and Ancient Egypt, and comes down to the 1960's. It is divided into three main sections: pre-Freudian, Freudian, and post-Freudian. Adequate accounts are given of the analytic dissenters, and there are critical comments on the rigidity of psychoanalytic doctrine. Of the three approaches to mental illness, magical, organic and pyschological, which the authors distinguish, they clearly believe the psychological is the most important, and it is for this reason that they regard Freud's work of inestimable value since "he succeeded for the first time in explaining behaviour in psychological terms." However, they believe that the future lies with a multidisciplinary approach in which psychoanalysis, organic medicine, psychology and sociology are integrated together.

The book is packed with information on every phase of psychiatric history and note is taken of relevant developments in medicine, general science, and psychology. The post-Freudian section is a particularly useful survey of the post-war scene in psychiatry which refers not only to neo-Freudian developments, but to organic methods, pharmacology, personality theory, social psychiatry, psychosomatics, the existential movement, and child psychiatry. American and central European psychiatry is certainly in this period much better represented than British or French work. Very few contemporary British psychiatrists receive any mention. Exceptions are Fairbairn, Maxwell Jones, T. F. Maine and Bowlby. It is also noticeable that the few misprints tend to relate to British names (e.g., Sully is twice spelt Scully, and Flügel is spelt Fliegel).

The book contains a number of challenging judgments and evaluations. The role given to Cicero as a founder of psychotherapy and a man gifted with exceptional clarity of thought is surprising. It is also startling to learn that in modern times "the role of the devil has been taken over by brain chemistry." A special appendix devoted to Jung's flirtations with National Socialism is a salutary reminder of events that are easily forgotten.

In general, the scholarly standard of the book is high, and it can be recommended, for its breadth and balance, as a most valuable survey both for the psychiatrist and the psychologist.

L. S. HEARNSHAW.

APTER, M. (1967). *An Introduction to Psychology.* Bristol: Teaching Programmes, pp. iv+637, 25s.

Dr. Apter is a stalwart of the tough-minded brigade and presents his own empirical evidence for the quality of this programmed introductory textbook: validation studies on first-year undergraduate and other students which show a substantial post-test gain on a criterion test and a low linear error rate on the programme. Unfortunately, in presenting this evidence, Dr. Apter falls from positivist grace by failing to state the numbers in his samples. The reviewer, whether he be tough-minded or otherwise, must therefore remain sceptical, read on and form his own judgment.

Introduction to Psychology, which contains chapters on Observation, Experimentation, Statistics, Learning, Perception, Individual Differences and Theorisation, is attractively produced and its programmed sequence is easy to follow. It is eclectic in its use of programming techniques and the style of writing, though rarely graceful and occasionally barbarous, is generally lucid. Dr. Apter has decided, wisely, to stress methodological issues; wisely, but not too well. As he admits in the Postscript, he has " deliberately avoided the discussion of a number of philosophical questions about scientific method." One might wish that he had been less reticent as the consequent portrayal of scientific method in psychology is frequently superficial, and, in parts, a gross caricature. For example, the operational definition of ' length ' favoured by Dr. Apter is " that which can be measured by using a ruler " (pp. 74-76). (With a little ingenuity, one could use a ruler to measure, among other variables, weight, reaction time and the specific gravity of certain woods). Or, " a result is significant when it is not due to chance alone " (p. 205). Also, in spite of an extended discussion on the importance of distinguishing between external and internal variables, Dr. Apter himself uses " sound (p. 111)," " pitch, timbre and loudness (p. 341) " when referring to the physical correlates of these variables.

The substantive content of this book is very thin indeed. The chapter on Learning, for example, presents only the rudiments of classical and operant conditioning; that on Perception covers Weber's Law, the Müller-Lyer illusion and a perfunctory treatment of the Gestalt School. Even so, errors of fact and misleading statements are not difficult to discover. It is not true that " *any* (Dr. Apter's italics) kind of stimulus can become a conditioned stimulus " (p. 533) or that a stimulus which is more intense than the C.S. elicits a C.R. of diminished strength (p. 341). The reader may also be misled on, among other things, the definition of an aptitude test (p. 448), the interpretation of the M.M.P.I. (p. 492) and Hull's concept of reinforcement (p. 606).

In view of its scantiness and its other inadequacies, *Introduction to Psychology* would not be suitable as a text in universities and colleges of education. It is difficult to take seriously Dr. Apter's suggestion that after reading it, students would be in a position to read research papers and undertake their own research.

BRYAN BETT.

BEARD, RUTH M. (1967). *Research into Teaching Methods in Higher Education.* London: Society for Research into Higher Education, pp. 51, 12s. 6d.

The Society for Research into Higher Education was formed three years ago to maintain the interest in research in this field which had been stimulated by the Robbins and Hale Reports. It has made good progress in preparing the way by showing how much scattered and heterogeneous research work is already being done in universities and colleges and by collating the findings of research already available. Roy Cox's survey of research into examinations in universities has been well received and now we have a survey on research into teaching methods in universities which promises to be equally useful.

This booklet, as its title indicates, is a survey of research, not a general treatise on teaching methods. The general reader would find it jerky and cryptic, but it is not meant for him. It is meant for researchers themselves and for others who want an operational map of this particular research area, either as a background for their own research plans or to ascertain what empirical evidence, if any, exists for the views on teaching which university people advance from time to time. For such readers the booklet is admirably compiled, grouping topics and giving the gist of relevant articles in a sentence or two for each.

Apart from details, one is left with four clear impressions. First, much is being done. The amount of interest shown by universities in their own teaching is greater than their critics suspect (though, as Dr. Beard says, the normal response of university teachers to a high failure rate is still to criticise either the examination or the original selection of the students). Second, much is *not* being done. In many fields we are told that there is little experimental work to report, at least in Britain.

Examples are: study skills, manual skills, unsupervised discussion groups, learning to make decisions, learning to think creatively in science, developing student attitudes, assessing complete courses (and also, though not mentioned by Dr. Beard, organizing laboratory work). Third, the diversity of learning situations in universities is such that all generalizations are dangerous: it is not easy to isolate the common elements when one student is listening to a lecture on Chaucer and another is learning how to fill a tooth. Fourth, the exercise of defining objectives, which research is increasingly forcing on university teachers, is a salutary one, whether or not any valid findings emerge in the end. STANLEY NISBET.

BIGGS, J. B. (1967). *Mathematics and the Conditions of Learning: A Study of Arithmetic in the Primary School.* London: N.F.E.R., pp. xi+441, 55s.

It is a thought-provoking task for one who is actively concerned with educational research to review this book. Dr. Bigg's main object was to compare the effectiveness of different methods of teaching arithmetic in primary schools. Given this aim it is hard to see how the design and execution of the research could have been improved. A sensible classification of teaching methods into traditional, structural and motivational was adopted and 5,000 children in eighty-seven schools were tested during their third junior year in 1959. Tests of attainment in arithmetical computation, problem-solving, the understanding of basic arithmetical concepts, anxieties about and attitudes to arithmetic, and verbal reasoning were employed. Asked to design an experiment in 1959, most competent research workers would have proceeded like Dr. Biggs—indeed imagination was shown in the decision to employ concept tests and measures of anxiety. Yet the results seem out-of-date, old-fashioned and almost irrelevant for the schools of 1968. Biggs is himself well aware of this for he writes in his preface: "This is a rapidly changing field, and the problems defined and investigated seven years ago may not be exactly those that are to the fore of the mind of the present reader . . . he may well feel that he is being tipped the winner of last year's Grand National." And in Chapter 10 "The whole field of mathematics education is in an ever-increasing state of flux. New techniques, new concepts, new ways of thinking about the subject, new slants on ultimate objectives and new insights on the learning process are born every day. This should be borne in mind in evaluating the present results. Even if the latter accurately depict the state of affairs existing in 1959, when the testing was carried out, are they really relevant to current problems in mathematical education? Further, can they provide a reasonable basis for attempting to meet problems for next year, for the next ten years? Research in any field is usually out of date to some extent before it is published; in the present case, this would seem to be more than usually true."

Exactly. A reviewer can only agree with these sentiments, applaud the author for his candour, and go on to speculate how many readers of the reports of the research in the educational press will also appreciate the sensible reservations of the author. What could be done to overcome these difficulties? More modest samples, less sophisticated experimental techniques, wider and more imaginative assessing and testing, quicker reporting are suggestions, though each has its own danger. The resulting research might be less scientific, less objective, but more relevant.

However, Dr. Biggs may be too modest. He is a competent psychologist who has an excellent acquaintance with modern teaching techniques in mathematics. Psychologists will find his analysis of current teaching methods of great interest, especially the description of the psychological principles involved in the Sterne, Cuisinaire and Dienes structural apparatus. Likewise, the chapter on "The Nature of Educative Learning," shows Biggs at his best and allows him to integrate his knowledge of psychology and mathematics in an illuminating fashion. His attempt to outline a model of the learning process which will have educational significance is a very fine one. The book is worth buying for this alone. Educational psychologists who lecture to teachers or potential teachers and who search, sometimes almost in vain, for examples of learning theory which have classroom relevance will find valuable examples here. JACK WRIGLEY.

Blackham, G. J. (1967). *The Deviant Child in the Classroom.* California: Wandsworth Publishing Co., pp. xii+191, $6.60.

This pleasantly produced book consists of twelve short chapters; within each chapter the material is clearly organised by headings and subheadings. The author, who is at Arizona State University, writes that this book is " addressed to teachers—in training and in practice . . . and others interested in children with a problem." He attempts to guide this class of reader to an understanding of deviant children in school by providing an outline of normal personality growth, an account of the varieties of parent-child relationships and a description of deviant patterns of behaviour. He also devotes a large part of the work to a consideration of the effects which the organisation of schools and the personality of teachers may have on problem children. He suggests many ways in which a teacher might gain insight into 'the deviant child' by direct observation and by talking to his parents. 'Deviance' is largely synonymous with the term 'emotional maladjustment' as used in the U.K.

The work contains much that is sensible, humane and insightful, communicated in terms that teachers, educational administrators and others can understand. It is a pity, therefore, that psychologists must be reluctant to recommend its use. The author's classification of maladaptive behaviour is psychiatric medical and there is no hint that this taxonomy is psychologically unsatisfactory. His account of the basic psychological processes and stages of mental growth is simply and uncritically psychoanalytic. No consideration is given to models, other than psychoanalytic, of intrapersonal conflict-induced behaviour, nor to the learning processes which maintain disturbed behaviour, nor to the part played by modelling and social imitation. There are also some surprising lacunae which are not made good by the references at the end of each chapter; (i) a few pages deal with brain-injured children as a deviant group, but there is no reference to the educational work of W. M. Cruickshank in this field; (ii) a section on 'self-concept' fails to direct the reader to Gordon Allport; (iii) no use is made of the many studies, such as the California longitudinal research, of the occurrence of behaviour disorders in 'normal' children.

The study of human behaviour and development is as difficult and complex as any scientific endeavour. Some ground has been won and most of us want to see the gains used by our fellows, especially such as school teachers. This book attempts to pass on these gains partly by practical suggestions, the fruit of the author's clinical experience, bearing on the conduct of teachers in classrooms, and partly by providing a conceptual psychological framework. The concepts offered are in the nature of a mythology of the mind—terms such as 'ego' and 'unconscious' are flagrantly reified—and they give poor service to the interesting case illustrations of childhood problems.

C. J. Phillips.

Crothers, E., and Suppes, P. (1967). *Experiments in Second Language Learning.* New York and London: Academic Press, pp. ix+374, £5 8s.

One of the outstanding failures of human experimental psychology has been the inability, in spite of many thousands of experiments in human learning, to produce results with direct application to classroom learning. This is a 'failure' not in the sense that it is the psychologist's duty to be useful, but that in striving for scientific purity many investigators have limited themselves to a very small, and not necessarily typical part of our memoric ability. In the past ten years or so linguists have discovered methods of describing languages which reduce somewhat the awesome complexity which drove many experimenters to use nonsense material in their enquiries. In addition, a new wave of eclecticism is gathering which might reduce the proportion of effort which has previously been devoted to vacuous controversy. Crothers and Suppes exemplify the potential of the future. Their theoretical base is stimulus-sampling theory, a variant of conditioning theory, but they cheerfully admit that the theory as it stands will be incapable in principle of accounting for all the facts of language learning. However, they point out quite correctly that the same is true for any other theory as it stands, linguistic or psychological, and provide a thoughtful discussion of the relation between psychological and linguistic theory which is distinguished by the absence of the authoritarian tone which mars most similar discussions.

After introducing the subject they proceed to describe a series of experiments which, as they admit in the preface, have only suggestive, not direct, implications for classroom learning. The experiments which were carried out appear to have been chosen with multiple motivation : general psychological interest, specific testing of hypotheses derived from models of the learning process, and relevance to the pedagogical problem. All the materials were from the Russian language, which means on the one hand that any positive results may be applied directly without need for assumptions of generality. On the other hand, the use of natural language material gives the reported work a sense of immediate applicability which is partly illusory. The publishers claim that the book will be of interest to language teachers, but most of it is so highly technical that only the most devoted teachers are likely to persevere long enough to derive much value, and the rest are likely never to read another psychological text.

Nevertheless, the book is virtually required reading for psychologists in the area of learning and psycholinguistics, and there are a number of observations that a linguist would find of interest. The sixteen experiments reported cover a range of topics from the discrimination of Russian phonemes to the inductive learning of Russian grammar. None of their conclusions is overwhelming in its significance, but throughout the text are scattered a large number of minor results for which a place will have to be found eventually in any model of language learning. Also, these results indicate further experiments likely to have direct application, and the models they develop appear to have considerable generality.

This book has the feel of being the first in a series. The next book will maybe combine data from further experiments based on the present work with practical results from the project in teaching Russian by computer currently in progress at Stanford University. That book should represent a real breakthrough on all fronts.

JOHN MORTON.

BOLLES, R. C. (1967). *Theory of Motivation.* New York : Harper and Row, pp. 546, no price.

There can be no doubt that this book constitutes a most comprehensive study of motivation, although it concentrates on the learning theory approach. The brief historical survey of motivational concepts is not only lucid but succeeds in setting the background for the ideas considered in more detail later in the book, illustrating, as it were, the sowing of the dragon's teeth. Freud, Lewin, Murray, McDougall and the ethologists are then each briefly described and the orthodox criticisms of the inadequacy of their model building made by way of introduction to the main corpus of the book, the section by which it must stand or fall—the analysis of learning theory applied to motivation. As is clear from the fact that there are around two thousand references, including much recent work, cited in the text this is a remarkably detailed study of the problems in this area and it is pleasing to note that no immoderate claims are made concerning the solution of these problems. Indeed, Bolles might, perhaps, appear to err on the side of caution. For example, " Such explanations (behaviour in terms of secondary reinforcement) put upon the theorist or researcher the very healthy responsibility of having to indicate precisely what stimuli, responses and reinforcements are involved in the behaviour being explained." This book, according to the author, is intended " for all those who would do research to improve our understanding of behaviour." This implies an advanced text which presumes a complete familiarity with the style, language and methods of psychological research together with a good grasp of the basic material. This is correct. It is not an easy book to read : each chapter needs and deserves careful consideration. Although a mine of references, it is far more than a book for reference since there is a steadily developing argument throughout the work. It would, therefore, appear to be a most useful book for students specialising in psychology. For those, however, who are interested in the more applied side, such as teachers or social workers, this would not be a suitable text, being too detailed, too concerned with rats and assuming too much knowledge of the subject. For its requisite audience a most valuable book.

PAUL KLINE.

FAIRWEATHER, G. W. (1967). *Methods for Experimental Social Innovation.* New York: Wiley, pp. 250, 60s.

The theme of this book is the application of scientific method to the study of social change. What emerges is a general discussion about the choice of problems for research, the formation of a research team, the collection and publication of data, and much else. At times there is sound (if unoriginal) advice. Thus, if you are interested in delinquency you are advised to start by living and working among delinquents; if you are the head of a research team you should invite the research workers to take part in the planning so as to increase their ' ego-involvement.' The advice, however, is often couched in such general terms as to be virtually without content. Thus, " Should any researcher's interest or behaviour be detrimental to the research effort, a determination can be made prior to the action phase of the research and termination of his association can be completed before the experimental subsystems are placed in operation " (pp. 63-4). In plain words this means (I think): " If a person is a nuisance, get rid of him early "; but it tells us nothing at all about how to deal with a *particular* misfit. This indeed, is the difficulty with the book as a whole. A few pages of original work describing how the author solved a *particular* research problem would, to me at least, have been of much more interest than many chapters purporting to be about ' research in general.'

T. R. MILES.

GOWAN, J. C., DEMOS, G. D., and TORRANCE, E. P. (Eds.) (1967). *Creativity: Its Educational Implications.* New York: Wiley, pp. xii+336, 64s.

This is a collection of 36 published papers drawn mainly from the *Gifted Child Quarterly*, which have been brought together as ' a first attempt to help educators transform the muddy beginnings of creativity into a flowing stream in the children they teach and counsel." This didactic intention is the least satisfactory aspect of the whole book. The British teacher would, I think, expect some justification for this intention, but one-quarter of the papers in the book are not accompanied by references and although six of the eight subsections of the book are followed by annotated bibliographies, that following the section on ' Teachers and Parents ' is very short and 'A curriculum for creativity ' has none at all. There is, of course, very little hard research to fall back on in this area, but it is less excusable that the section, ' Guidance and Measurement of Creativity ' should again lack an annotated bibliography or that two of its constituent papers should be without references.

This leads to a general impression of consensus, no doubt essential to the didactic purpose; there are no more than three or four fleeting references to the creativity/intelligence distinction—the lack of an index makes them difficult to track down—and none of these is of a type to suggest to the naive reader that this is a controversy. The author index further reveals that of the 36 references to Getzels and Jackson's book there is no overt mention of any critical comment apart from one cryptic instruction, ' (Cf. de Mille and Merrifield, 1962).'

In spite of its declared intention, this is not a book to recommend to students as a basic text on creativity but there are half a dozen review articles which by virtue of their evaluation of the field in general or their bibliography of research on the educational variable in creativity should earn it a place in a college library.

T. CHRISTIE.

JAKOBOVITS, L. A., and MIRON, M. S. (Eds.) (1967). *Readings in the Psychology of Language.* New Jersey: Prentice-Hall, pp. xi+636, 80s.

Though expensive, this book is very good value for anyone interested in the psychological analysis of linguistic behaviour. The editors have not only succeeded in bringing together some of the most important papers on psycholinguistics written during the last decade, but they have marshalled their material in such a way as to reveal very clearly the nature of major controversies in this rapidly expanding field of study. The deliberate placing in contiguity of opposing viewpoints can lead to confusion for the reader, but in this book the selection of writings is so skilfully done that each topic is simply seen in greater depth. Thus, we have the views of B. F. Skinner on "Thinking" (taken from his book *Verbal Behaviour*, 1957) followed by Chomsky's critical review of the book which, in dealing with Skinner's contentions, also points to the weaknesses inherent in the behaviouristic approach to higher mental processes generally. Or, a particularly praiseworthy inclusion this, there is available the whole sequence of the argument between M. D. S. Braine and Bever, Fodor and Weksel, which originated in Braine's paper "On Learning the Grammatical Order of Words," and which crystallises the differences between contemporary linguistic analysis and current learning theory.

Of course, not all the papers have the intended relationship of these two examples. In other cases the complement arises from the editors' careful arrangement, or the view in depth is preserved by their selection of representative papers which trace a recent history of ideas on the topic (e.g., C. N. Cofer's "On Some Factors in the Organizational Characteristics of Free Recall."). The result is a first-rate book which must be of genuine help to the advanced student, lecturer and research-worker. Of the three sections into which the book is divided—the theroetical development of psycholinguistics, experimental approaches to language, and the problem of meaning—it is probably the last which will make the most immediate appeal to the interests of workers in the field of education. It contains, amongst other good things, J. B. Carroll's article, on "Words, Meanings and Concepts" which forms a useful bridge between the psychology of language and concept formation and practical teaching problems. With the formidable growth of the experimental science of language, such reflective attempts at application are all too rare, and the research-minded teacher who wishes to tackle this book should make Carroll's paper his first selection for reading.

D. J. BRUCE.

NASH, P. (1966). *Authority and Freedom in Education—An Introduction to the Philosophy of Education.* New York: Wiley, pp. 350, 53s.

This volume is not an introduction to the philosophy of education as such but (as the author admits) to the *theory* of education. In other words, although it is concerned with "... the clarification of terms and meanings and the substantiation of arguments ...", it mainly involves the drawing together of relevant philosophical, psychological and sociological findings in the formulation of educational judgments and prescriptions. In order to achieve this, the author has divided the major topic into a number of 'sub-polarities' such as 'The Authority of Discipline/The Freedom to Develop One's Interests,' or 'The Authority of Determinism/The Freedom to Choose.' With each, he employes analyses and data of various kinds in a dialectical framework which ultimately issues in a series of synthetic educational prescriptions embodying his own commitment. Thus, after an examination of the notions of 'group' and 'individual,' he reaches an inclusive resolution of the tension between them in the form of various recommendations for fostering the education of autonomous individuals committed to serving the community. The remaining 'sub-polarities' are dealt with in similar fashion.

The book is interesting because of this and should prove a useful addition to any education library. However, it has one or two shortcomings which require mention. The author displays his own commitment very openly and urges his readers to do the same, in the belief that students will best learn the enterprise of theorising about education " . . . by taking a sincere stand on an important issue and learning its justification and weaknesses . . . " So far, so good—but surely such commitment must be *argued* (rather than asserted) if it is to have the desired pedagogical effect? Again, psychologists (and others) may boggle occasionally at the author's use of extra-philosophical data (" Infants deprived of love tend to become morally and socially defective " or " The juvenile delinquent is a model of adjustment to his own delinquent group," or " What the theologian calls sin produces . . . what the psychiatrist calls neurosis or psychosis "). H. M. PATERSON.

PASSOW, A. H., *et al.* (Eds.) (1967). *Education of the Disadvantaged—A Book of Readings*. New York: Holt, Rinehart and Winston, pp. 503, 60s.

At last a book within the hazy boundaries of 'Education' that is not merely a collection of exhortations and woolly laments! The toughest-minded psychologist can learn from this book, and may become impelled to widen his research horizons, because it is a practical and eclectic application of both Psychology and Sociology to the frighteningly obstinate problems of educating the economically and (therefore) culturally impoverished.

The readings are skilfully woven from current research in teaching experience and from accounts of special programmes in depressed areas. But there are also some papers of more general psychological significance than the narrow limits of " school achievement." Bernstein's work on social class and language is reported; J. McV. Hunt discusses the psychological basis for the enrichment of early school-experiences to stimulate later cognitive and affective development, and criticises cogently " the belief in fixed intelligence "—that still appears from time to time in educational circles. Ausubel, in a tautly-argued paper, discusses the reversibility of the effects of cultural deprivation, and the development of motivations for academic achievement. He observes that in intellectual development there are " optimal periods of readiness for all kinds of cognitive development," and that, as in stages of physical development, we may encourage or discourage growth less grossly and more long-lastingly than has hitherto been suspected.

The book is divided into five sections: (1) The meaning of the term 'educational disadvantage' is discussed. (2) The specific and differing patterns of cognitive and motivational drawback of the minority-groups in the U.S.A. are described, and related to the nature of their educational backwardness. Zintz's study of American-Indians suggests that 'integration' *per se* is no panacea for cultural deprivation. What if the group, 'deprived' in terms of the major culture, has its own strength and integrity and motivations? (3) The social-psychological factors influencing school-achievement are discussed, and this section usefully deals with the cognitive and linguistic preparation for schooling. Possibly more reference might have been made to 'pure' psychologists such as D. E. Berlyne, and certainly more reference should have been made to the growing number of linguists concerned with the teaching of a *second* language in a multi-lingual society, such as South Africa. (4) and (5) are the practical core: what is being done for the disadvantaged pupil, and what sort of teachers can best assist. A recurring theme is the need to bridge the gap between the middle-class teacher and the lower-class pupil.

The book is disturbing: how many 'pure' psychologists coyly ignore the educational problems which are both practical *and* of theoretical importance? Those of us who are trained in the more narrowly conceived areas of social and experimental psychology could do themselves a professional service—and benefit their society—by applying their precise techniques to the topics of this book.

The next edition of the book ought to be less limited to the U.S.A. The problems of immigrants in the U.K. who speak little English and the countries in Africa, Asia and Latin America faced with mass-ignorance, semi-literacy and barred (by their history) from the mainstream of development, make this type of book deplorably pertinent for the remainder of the twentieth century. LEONARD BLOOM.

PETERS, M. L. (1967). *Spelling: Caught or Taught.* London: Routledge and Kegan Paul, pp. viii+88, 15s.

Mrs. Peters has provided a very useful addition to this helpful series. Since the old drill methods of teaching spelling have attracted the suspicion attached to formal linguistic teaching, many teachers have "comforted themselves with the doctrine that spelling is picked up incidentally," through interesting reading and creative writing. The book is a thorough and very compactly written survey of research which makes it clear that spelling has to be taught.

She argues cogently the case for good spelling and then proceeds to analyse the nature of the skills involved in terms of research work on sensation, perception, imagery, motivation and self-confidence in the learner. A survey of traditional methods emphasises that spelling-lists should be concerned not only with what children can spell at various stages of development but with what they want to spell, and that the lists must be supplemented by specific teaching of the children in the best methods of learning, involving oral, written and kinaesthetic approaches—Fernald's "amalgam of the sensory output." Much of this is practised by effective teachers using more traditional methods, but teachers using the less personal approaches demanded by creative writing will find stimulating suggestions in the subsequent discussion of current approaches which make intelligent use of the children's activity and verbal environment, not excluding T.V. advertisements which are not only visual and oral but also emphatic, repetitive and strongly interesting to children. There are also stimulating discussions of the assessment of spelling and the precise points to which the learning of children is to be directed.

Mrs. Peters is scholarly and precise and many teachers will find her book demanding, but rewarding; students in training will find in it a sound basis for discussion and experiment, and for advanced students there is a very good bibliography.

A. H. STEWART.

RAVEN, J. C. (1966). *Psychological Principles Appropriate to Social and Clinical Problems.* London: H. K. Lewis, pp. x+187, 25s.

This is a highly individual book, relatively short though not all that easily read because of its rather diffuse abstract style. It is prefaced by a prologue which takes the form of an edited transcript of a tape recorded conversation with the author which adds little to the burden of the text other than a passing but succinct statement of the aim of the book. Raven says, "I have tried to describe how I think about behaviour and how I think we can compare one form of behaviour with another." By the end of the book, however, he has become more specific and considers that the four questions embodying the principles he has stated, which the psychologist must answer in respect of a patient are: "(1) How does a person's attention vary with time in its location, range and concentration? (2) Does his behaviour show well defined intentions with respect to the past and future? (3) What does he enjoy as good or fear as evil? (4) How does he order his thinking with respect to himself and his surroundings?" He has "tried to show how along these four variables or 'co-ordinates' of any and every kind of interview, the psychologist is able to demonstrate characteristic differences in the things people say or do." He has "tried to show the extent to which he can at present assess these four variables in a person's behaviour so that they can be correlated with a person's organic constitution and with the social factors influencing his actions as assessed by other members of the team," and he has "also indicated problems which await psychological research."

Between these two main statements fall five chapters under the general heading of Practical Problems followed by eight further chapters on Theoretical Principles. In these chapters Raven's first aim has been reasonably effectively met but his secondary aims perhaps less effectively met. All of them reflect a rather mentalistic, even anti-behaviourist, view in which philosophy rather than biology is seen as the proper bed-fellow of modern psychology, but within this framework the general theory propounded hangs together reasonably well. The trouble with it is that it gives a conceptual structure for the process of behaviour assessment only, and has little to offer the business of behaviour prediction or mediation.

In the first chapters particularly, there are several irritating naiveties, as when Raven's arguments appear to presuppose the veridical rather than the operational nature of questionnaire responses in the work of Eysenck and Cattell (pp. 45-48); and when he says, "As soon as we become interested in what a person means by the words he uses, our investigations must necessarily be more subjective" (pp. 10-11). One feels that Osgood and Kelly might feel bound to disagree. Or on p. 88, "It is this decision to respond or not to respond in some habitual way to a given stimulus which is characteristic of *mental* as distinct from mechanical activity." What decision? It is, after all, only inferred to have occurred.

Later, as the book develops its theme—and the relevance of planning and the experience of intention to the growth of the self-image is well pointed—there is more of value to be found and a second reading certainly helps. One can forgive two names misspelt in the bibliography and several textual references (pp. 44 and 47) omitted from it, if only for Raven's endorsement of the view that "It is from the time a person begins to realize that he himself plays a part in the assessment he makes of another person's behaviour that he really becomes a psychologist" (p. 161).

For those who have the patience to read it twice, the book is just worth the money, but it is a book for the armchair rather than the laboratory. D. F. CLARK.

RICHARDSON, E. (1967). *Group Study for Teachers.* London: Routledge and Kegan Paul, pp. 130, 15s.

The aim of this book is to help students and teachers to understand something of the emotional life of classroom groups, of which pupil-teacher relationships are a very important part. While such an aim is highly commendable, in practice there can be little doubt that its realisation is difficult. Hence any contribution directed towards this end is worth serious consideration.

Contrary to what might be expected from the title, this book does not present systematic treatment of methods which teachers might use to study groups. Rather is it an account of the results of one method which has been used by the author, a tutor in a university department of education. This is a method in which students, through membership of a small group, study the dynamics of group behaviour from personal experience. Such study is done under the leadership of a consultant, in the manner of the Leicester-Tavistock conferences, which the author has attended in such a capacity.

The material which the author has used is drawn from six years' experience with small student groups—voluntary study groups, tutorial and seminar groups. Her method is perhaps best described in her own words "I can produce nothing from these experiences that is in any way quantitative or amenable to statistical analysis. I never made notes during a tutorial or study-group meeting: and although there was a time when I tape recorded all discussions in certain groups, with their knowledge, I eventually stopped doing this, partly because I felt that the recording machinery was bound to have some effect on what was happening, and partly because I found the transcripts so voluminous and so unselective that they were virtually useless. All my evidence, then, is in a very real sense subjective. My memories of the meetings were recorded as soon as possible after they took place." This approach is open to considerable methodological criticism, but it is important to bear in mind that this book is a pioneer endeavour and that, presumably, in the future, impressionistic methods of study will be combined with more systematic observational techniques.

This subjective material is used by the author in relation to a number of problems which she has observed in her contacts with students in the Department of Education and which she feels are also found in school classroom groups—the social consequences of evaluation being part of the teacher's role, the effects on the group of the teacher adopting a research role, the importance of furniture to group behaviour, the social significance of lateness and withdrawal, and the termination of group relationships. It could be argued that, in a book for students who would not have the group experiences which her own students had, material gathered from classroom situations would have been more relevant to the basic aim. An earlier book by the same author in fact provides this.

The discussion of these problems is carried on largely in terms of Bion's theory of group behaviour, which is given extremely brief treatment in the first chapter. Such brevity might be unavoidable in such a slim volume, but the interpretation of selected incidents in school groups without reference either to significant experiences that may have influenced the behaviour of individuals, or to the structural or normative characteristics of the group within which the behaviour is taking place, reflects the necessarily limited perspective of the book. The author is aware of the dangers associated with this method, but considers the gains in understanding which follow from membership of study groups to be well worth such risks. The gains which follow from reading about discussions which have taken place in other groups she regards as less valuable than actual participation.

As an introductory text for students this book could be useful, so long as additional reading was suggested. The way in which it is used, however, is obvioulsy of considerable importance.

DOUGLAS S. FINLAYSON.

RYLE, A. (1967). *Neurosis in the Ordinary Family.* London: Tavistock, pp. 156, 32s. 6d.

Dr. Ryle is a general practitioner already well known for his interest in psychiatric problems and the epidemiological approach. The survey on which he now reports concerned a group of 112 families in his National Health practice. The work was on a team basis, but the part of the psychiatric social worker, Miss Madge Hamilton, who collected most of the data, was particularly important.

The population involved consisted entirely of working class families with children of primary school age. The degree of neurosis present in the parents was assessed by using the Cornell Medical Index and by other means. The children's adjustment was judged from teachers' reports (on Douglas and Mulligan's Schedule) and interview information based on the Macfarlane Scale. The results showed that 14 per cent of fathers, 35 per cent of mothers and 28 per cent of children were in the neurotic category. Further findings, again mainly derived from approximately four hours of interviewing each family, indicated that neurosis in the parents was related to their own unhappy or insecure childhood experiences, was connected with marital failure when this occurred, and had its effect in developing poor attitudes towards the children, who themselves tended to be maladjusted.

The book is interesting and embodies unique material. It is clearly written and some well-chosen case histories help to bring the realities home. Some of the procedures, as Dr. Ryle admits, are not completely satisfactory and it would certainly be unwise to generalise extensively from the results. These are, nevertheless, highly suggestive and deserve the serious consideration of medical administrators, general practitioners, and all social workers.

O. C. SAMPSON.

SPENCE, K. W., and SPENCE, J. T. (1967). *The Psychology of Learning and Motivation.* Vol. 1. New York: Academic Press, pp. x+381, 92s.

This volume is the first of a series under the same general rubric. This rubric is intended to cover work in such fields as classical and instrumental conditioning in human and animal subjects, complex learning, memory and problem solving activity, and studies of primary and acquired motivation. The editors propose that the aim of the series is to provide a vehicle in which psychologists specialising in the field might attempt an integration of research data either at the empirical level or "orientated around the evaluation or development of some theory." This aim flows from a desire to make possible a systematising of some of the research findings which have followed the work of the earlier system builders. To this end each volume is to contain a number of contributions by researchers in the field, which will be more substantial statements than is normally the case in the psychological journals.

The present volume has five such contributions. Three are concerned with aspects of learning and reinforcement in animals (predominantly the rat). Two are concerned with human learning. The first three comprise a paper on partial reinforcement by Amsel, one presenting a sequential hypothesis of instrumental learning by Capaldi, and one on satiation and curiosity by Fowler. These three papers derive a good deal of their strength from the fact that the authors are able to review specific aspects of the field and to place it into a context of their own research.

The present reviewer found the contribution of Fowler of particular interest in view of his bid to integrate the ideas of appetitive and exploratory motivation. To educationalists, however, these three contributions may well be of peripheral interest. The remaining two are more relevant. In the first one, Bower argues for a multicomponent theory of the memory trace, attempting, as he says, to " demonstrate a common theme running throughout diverse branches of the research tree of human memory." To this end he examines such aspects of human memory as encoding, recall and recognition. This contribution is mainly an exposition of a mathematical model, but it draws on some ingenious experiments.

Of the collection of papers, the one by Mandler on organisation and memory will probably be of most interest to educational psychologists. In this paper Mandler reports a number of experiments which take as a point of departure Miller's unitization hypothesis on the processing of information. The experiments explore the relationships between categorisation and recall of verbal units. Mandler uses memory as a way of evaluating organisation or structure in information processing, and his experiments provide evidence for a hierarchical view of cognitive structure which chimes well with Miller's views and those advanced by Ausubel and Bruner. Finally, he suggests that future investigations on the development of organisation should focus essentially upon language development and semantics. The present reviewer found this contribution of particular value since it adduced new experimental evidence in a field of great relevance to pedagogy through its rigorous approach to the question of 'structure' which, as Mandler remarks, has a somewhat tattered reputation.

E. Stones.

Stephens, J. M. (1967). *The Process of Schooling*. New York : Holt, Rinehart and Winston, pp. viii+168, $3.95.

In this eminently readable book, an attempt is made to outline a psychological theory of schooling. It seems a curious undertaking. I can see no more compelling reason for expecting that the causative influences at work in schools would be sufficiently unified, coherent, and specific to merit a separate theory, than for making a similar assumption with regard to, say, juvenile delinquency, or indeed any administratively-defined area of behaviour. Nor does the author convince me otherwise, although, in the course of his argument, he makes some valid, and perhaps neglected points.

Schools, it is argued, evolve because, although a child's parents can be relied upon to equip him with those behaviour patterns necessary for his immediate survival and their immediate convenience, there remain certain 'frivolous, decorative, or playful tendencies' of a manipulative or exploratory nature, which may have enormous long-term survival value for the group. These are unlikely to be of direct concern to the parents, and ensuring their cultivation requires a special agency, the school.

Now there are obviously other pressures which have worked towards the establishment of schools, ranging from urgently-felt needs to instruct children on matters with which parents have little or no competence, to the reluctance of mothers to have their social lives interrupted by children not yet of statutory school age. But although the author grants passing recognition to some of these in his argument, they are not incorporated in his theory.

A second area of discussion concerns the actual working of the schools. The process of schooling derives from certain spontaneous tendencies on the part of the teacher: to be interested in the less vital areas of experience, with which the school deals; to hold forth about his interests on little or nor provocation; to react with approval or disfavour to the utterances of others within his field of interest. These, not the deliberate efforts of the teacher, not the decisions of the administrator, are the heart of the schooling process, and supply (from the learner's point of view) the conditions for learning. Since these forces operate without our intervention, we are urged to take a relaxed attitude to certain issues of current concern. The content of the curriculum can (and indeed will) be determined by the interests of the teacher. The posited tendencies being prevalent there is an ample reservoir of adequate teachers; since they are spontaneous, training is no great problem. There is certainly little point in worrying about administrative or organisational matters.

It is salutary for those concerned with educational research or professional training courses to be reminded that teachers are human; that their behaviour is determined by motives more basic than the cerebral considerations we urge upon them. One might quarrel with the list of spontaneous tendencies invoked, and the verbal, expository view of teaching they imply. One might indicate the existence of other tendencies (that, for instance, to respond to aggression with counter-aggression) less likely to have beneficent results. But it is the issue of testability which raises the most serious doubts. Supporting evidence for the theory—the failure of most educational innovations to produce any lasting improvement—is of a purely negative nature. The author readily admits that non-significant results do not really advance anyone's argument, but I am inclined to think that the theory allows of no other kind of test.

W. B. HANTON.

WEBB, L. (1967). *Children with Special Needs in the Infants' School.* London: Colin Smythe, pp. 216, 30s.

Lesley Webb presents forty-four short studies of children who attended an infants' school in the South of England and were considered special problems. Children with similar problems are discussed in separate chapters dealing mainly with aggression, withdrawal, anxiety, pilfering, immaturity, bizarre behaviour, sex worries, physical limitations and cultural deprivation. The individual studies in each chapter are preceded by a brief definition and discussion of the behavioural problem and followed by a consideration of ways in which teachers, in co-operation with parents and others, can help children to cope more effectively.

The author rightly stresses that the teacher's primary concern should be in educating children and that it is not her job to diagnose causes or to seek explanations for aberrant behaviour, a task which requires the expertise of the psychiatrist or the psychologist. Her purpose in writing the book is to help teachers to be more effective in recognising children who need special help and in making decisions as to the most appropriate course of action to be taken. Throughout the book Miss Webb stresses that it is important for teachers to regard children as individuals with their own unique pattern of needs, rather than *en masse* in terms of generalised theoretical concepts. In the penultimate paragraph she warns the reader " that a course of action found efficacious with one child will not transfer to another, and an attempt to use courses described here as ' ready-made ' remedies to apparently similar cases will end in failure."

It is a pity that the quality of the material presented should be marred by so many printing errors which are dispersed throughout the book. Some statements are somewhat difficult to interpret. For example, " Between child and child, and within an individual, there will be anomalies of development which are natural and normal . . . " (p. 126). However, in spite of these defects, this book is a very useful contribution to the field of infant-education.

ROY HARRIS.

WOLFGANG, M. E., and FERRACUTI, F. (1967). *The Subculture of Violence: Towards an Integrated Theory in Criminology.* London: Tavistock, pp. xxiii+387, 63s.

This book argues the case for an 'integration' of different approaches (e.g., sociological, clinical) to the study of crime, and the idea of 'subcultures of violence' is advanced as an integrating concept in relation to violent crime, especially homicide. The book consists of five chapters. The first chapter discusses the meaning of integration and various possible senses of the term. The authors favour what they call 'fusion,' in which researchers belonging to different disciplines accept *one* theoretical system within which eclectic data collected by methods appropriate to the different disciplines may be handled. Chapter two is concerned with 'criminology as an integrating discipline.' The authors claim that we have so far had mainly *multidimensional* rather than inter-disciplinary studies, each discipline collecting facts within its *own* perspective. Sociological and psychological factors which lead to criminality tend to be treated as different, and separate processes; and we are relatively lacking in attempts to *relate* sociological to psychological factors.

In chapter three, an attempt is made to develop the concept of 'subculture of violence' as a concept encouraging such attempts at relation. It is suggested that the internalization of sub-cultural values (e.g., supporting violence) could be assessed by such methods as scaling of allegiance to these values and their implications. In the form of violence represented by homicide, the groups with the highest homicide rates should also show the greatest attachment to subcultural norms of violence. Chapter four considers the contribution of biological, psychiatric and psychometric studies to the study of homicide, and suggests that it is important to observe the difference between 'normal' and 'pathological' crime and to see *how* these fit together. Chapter five reviews social investigations of homicide, refers to various aspects of the differential distribution of homicide, and includes a brief account of extremes of 'cultural violence' as in Colombia and elsewhere. There is also in this chapter an extended section on treatment and control at various levels. The book is very fully documented, each chapter being followed by a long list of notes and references, and would be useful to many people for this reason alone.

'Integration' of different disciplines such as sociology and psychology is a thorny problem, and many people believe that psychological and sociological points of view represent different perspectives which can be used on the same empirical problems, but not combined. The case which the present authors make is for integration in criminology, in the sense that one system of concepts may generate and weld together both sociological and psychological research. The present reviewer, while impressed by this book in many ways, does not feel that the case has yet been sufficiently proven.

D. GRAHAM.

NEW JOURNALS

New Research in Education.
Scottish Educational Studies.
The Irish Journal of Education.

There could be few more convincing pieces of evidence for the upsurge of interest and activity in education than the recent proliferation of journals in this field. In this edition the editors are glad to welcome three new journals, each of which seems likely to make a distinctive contribution.

The current pressure for the rapid communication of research findings has led to criticism of some of the recent publications of the National Foundation for Educational Research as being partly out of date by the time they appeared. *New Research in Education* represents an attempt by the Foundation to provide a speedy outlet for interim research reports on long-term projects, mainly derived from the Foundation's own programme. Thus, in the first volume, dated June, 1967, nearly half the space is devoted to two interim reports by Joan Lunn on *The Effects of Streaming in the Primary School*. These provide a clear account of a thorough piece of research and should be a timely contribution to the current debate, unless their comparison of attainments in different types of school, however cautiously qualified, is too contrary to the prevailing educational climate to survive ! Of the remaining nine papers, that by B. S. Cane on *Ninety-one Good Readers* provides a very interesting supplement to the work reported by J. M. Morris in *Standards and Progress in Reading*.

This new journal will be published annually, or more frequently if necessary. Subscriptions (15s.) should be sent to Newnes Publishing Co., Ltd., Tower House, Southampton Street, London, W.C.2.

As its editor remarks, one remarkable thing about *Scottish Educational Studies* is that it should appear so belatedly in an age when nearly every English Institute of Education publishes some kind of journal of its own. Because most Scottish work has been published in national journals, it has been difficult to gather any impression of the separate nature and weight of the Scottish contribution. The new journal hopes to begin to remedy this and the first number is an attractive blend of articles of general interest such as that by Drever on *The Nurture of Intelligence* and those of particularly Scottish interest such as Elizabeth Perrott's article on *Stirling University's Plan for the Preparation of Teachers*. The final section, on ' Research in Progress ' provides an unusually wide and detailed sample of what is happening in Scottish education : this section could well be expanded in later issues.

Scottish Educational Studies is published by the Department of Education, Glasgow University, in association with the corresponding Departments in the other Scottish Universities. Contributions and subscriptions should be sent to The Editor, ' Scottish Educational Studies,' 4, University Gardens, The University, Glasgow, W.2.

The most recent journal is the *Irish Review of Education*, which is to be produced twice yearly, in Summer and Winter, from St. Patrick's College, Dublin, 9, to which address annual subscriptions (13s.) should be sent. This journal is the only one of the three aimed mainly at interesting the classroom teacher and helping him to keep abreast of developments in education, and this seems to be reflected in the papers, most of which manage to combine scholarship with lucidity of style and lightness of touch and to bring together a wide range of articles in an important area of education. Kellaghan's article on *The Organisation of Classes in the Primary School* is especially successful in this last aim and he even manages to include a reference to the previously mentioned article on streaming in the first number of *New Research in Education*. The journal is ambitious in its aim to include articles on any aspect of education and it should be helped in this by having secured the services of an impressive panel of consultants.

If these three journals can maintain the general quality of the articles which appear in their first editions, then each will deserve a wide circle of readers.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. " n.p." means " no price stated."

ADLAND, D. E. (1967). *The Group Approach to Drama—Pupils' Book Four.* London : Longmans, pp. 179, 7s. 6d.

ARGYLE, M. (1967). *Psychology and Social Problems.* (Paperback edition.) London : Tavistock ; New York : Barnes and Noble, pp. 232, 16s.

ARMSTRONG, D. (1968). *A Materialist Theory of the Mind.* London : Routledge and Kegan Paul, pp. 372, 50s.

BASSETT, G. W., *et al.* (1967). *Headmasters for Better Schools.* Queensland : University Press, pp. 109, $3.50.

BERKOWITZ, L. (Ed.) (1967). *Advances in Experimental and Social Psychology—Vol.* 3. New York : Academic Press, pp. 333, 84s.

BOLLES, R. C. (1967). *Theory of Motivation.* New York : Harper and Row, pp. 546, n.p.

CHASSAGNY, C. (1968). *La Lecture et L'Orthographie Chez L'Enfant.* Paris : Presses Universitaires de France, pp. 261, F12.

CLAUSSE, C. (1968). *Pedagogie Rationaliste.* Paris : Presses Universitaires de France, pp. 186, F9.

COFFEY, H., and WIENER, L. (1967). *Group Treatment of Autistic Children.* London : Prentice-Hall, pp. 132, 40s.

COHEN, J. (Ed.) (1968). *Psychology—An Outline for the Intending Student.* London : Routledge and Kegan Paul, pp. 203, 25s.

CROPLEY, A. (1968). *Creativity.* London : Longmans, pp. 121, 12s. 6d.

DAVIES, M. (1967). *The Use of the Jesness Inventory on a Sample of British Probationers.* London : H.M.S.O., pp. 20, 2s. 3d.

DECONCHY, C. (1968). *Structure Genetique de L'Idee de Dieu.* Brussels : Lumen Vitae, pp. 235, $8.50.

DINNAGE, R., and KELLMER PRINGLE, M. L. (1967). *Residential Child Care : Facts and Fallacies.* London : Longmans, pp. 344, 30s.

DINNAGE, R., and KELLMER PRINGLE, M. L. (1967). *Foster Home Care : Facts and Fallacies.* London : Longmans, pp. 268, 30s.

DUNN, S. S. (1967). *Measurement and Evaluation in the Secondary School.* Victoria : Australian Council for Educational Research, pp. 131, $3.00.

FELLOWS, B. J. (1968). *The Discrimination Process and Development.* Oxford : Pergamon, pp. 218, 75s.

FIELD, E. (1967). *A Validation Study of Hewitt and Jenkins' Hypothesis.* London : H.M.S.O., pp. 20, 2s. 9d.

GLEASON, G. (1968). *The Theory and Nature of Independent Learning.* Pennsylvania : International Textbook Co., pp. 101, 27s.

HERZBERG, F. (1967). *Work and the Nature of Man.* London : Staples Press, pp. 203, 25s.

HESS, R., and TORNEY, J. (1967). *The Development of Political Attitudes in Children.* Chicago : Aldine Publishing Co., pp. 288, $9,75.

HOLMES, M. (1967). *The Comprehensive School in Action.* London : Longmans, pp. 88, 10s. 6d.

HOOPER, D., and ROBERTS, J. (1967). *Disordered Lives.* London : Longmans, pp. 87, 10s. 6d.

ISAMBERT, A. (1968). *L'Education des Parents.* Paris : Presses Universitaires de France, pp. 245, F7.

ISMAIL, A., and GRUBER, J. (1967). *Motor Aptitude and Intellectual Performance.* Ohio : Merrill Books, pp. 199, $4.95.

JERKEDAL, A. (1967). *Top Management Education.* Stockholm : Swedish Council for Personnel Administration, pp. 237, C38.

JACOBSON, E. (1967). *Tension in Medicine.* Illinois : Charles C. Thomas, pp. 148 $7.75.

Kohnstamm, G. (1968). *Piaget's Analysis of Class Inclusion: Right or Wrong?* The Hague: Manton and Co., pp. 153, G15.

Komisar, B., and Macmillan, C. (Eds.) (1967). *Psychological Concepts in Education.* Chicago: Rand McNally, pp. 255, cloth 57s., paper 38s.

Lipsitt, L., and Spiker, C. (Eds.) (1967). *Advances in Child Development and Behaviour, Vol.* 3. New York: Academic Press, pp. 272, 76s.

Lorenz, K. (1968). *On Aggression.* London: Methuen Paperback, pp. 273, 10s. 6d.

Macdonald, J. (1967). *Careers Decisions Project—Basic Data.* Ottawa: Duhamel, pp. 203, n.p.

Marklund, A. and Soderberg, P. (1968). *The Swedish Comprehensive School.* London: Longmans, pp. 119, 15s.

Meyer, W. (1967). *Readings in the Psychology of Childhood and Adolescence.* Massachusetts: Blaisdell, pp. 346, $5.25.

Miller, G. (1968). *The Psychology of Communication.* London: Allan Lane, pp. 197, 35s.

Nickson, N. (1967). *Education through Music.* Queensland: University Press. pp. 35, 75c.

Parfit, J. (Ed.) (1967). *The Community's Children.* London: Longmans, pp. 108, 15s.

Parkin, G. W. (1967). *Success and Failure at the University. Volume II—The Problem of Failure.* Wellington: New Zealand: Council for Educational Research, pp. 233, $NZ 3.60.

Payne, D., and McMorris, R. (1967). *Educational and Psychological Measurement.* Massachusetts: Blaisdell Publishing Co., pp. 419, n.p.

Platt, Lord, and Parkes, A. (Eds.) (1967). *Social and Genetic Influences on Life and Death.* Edinburgh: Oliver and Boyd, pp. 222, 63s.

Rosenblith, J., and Allinsmith, W. (1967). *The Causes of Behaviour: Readings in Child Psychology and Educational Psychology.* Boston: Allyn and Bacon, pp. 608, 46s.

Stolz, L. M. (1967). *Influences on Parent Behaviour.* London: Tavistock, pp. 355, 65s.

Telford, C., and Sawrey, J. (1967). *The Psychology of Adjustment: Second edition.* Boston: Allyn and Bacon, pp. x+466, 57s.

Tolstoy, L. (1968). *Tolstoy on Education.* Chicago: University Press., pp. 360, 54s.

Wathen-Dunn, W. (Ed.) (1967). *Models for the Perception of Speech.* Massachusetts: The M.I.T. Press, pp. 470, 80s.

Watzka, L. (1968). *Kleinkind and Fernsehen.* Vienna: Osterreichischer Bundesverlag, pp. 180, DM26.

Williams, P., and Gruber, E. (1968). *Response to Special Schooling.* London: Longmans, pp. 160, 15s.

Younghusband, E. (1967). *Social Work and Social Values.* London: Allen and Unwin, pp. 200, 35s.

—— (1967). *Putting Research, Experimental and Demonstration Results to Use.* Washington, D.C.: U.S. Department of Labour, pp. 71, n.p.

—— (1968). *Language Teaching Abstracts, Vol.* 1. Cambridge: University Press, pp. 64, 20s.

—— (1967). *Register of Research into Higher Education.* London: Society for Research into Higher Education, 30s. non-members.

—— (1967). *Prospective* 14, *Education et Societe.* Paris: Presses Universitaires de France, pp. 203, F13.

IMPRESSIVE BOOKS ON PSYCHOLOGY FROM SAUNDERS

Esper :
A HISTORY OF PSYCHOLOGY
By Erwin A. Esper (Emeritus), University of Washington.
368 pages. £2 15s. 0d.
Published February, 1964.

Here is a superbly written history of psychology from its beginnings in early Greece to the investigations of the present day.

Hebb :
A TEXTBOOK OF PSYCHOLOGY (2nd edition)
By Donald Olding Hebb, McGill University
353 pages. Illustrated, £2 9s. 0d.
Published February, 1966.

This best selling text offers a revised and expanded clarification and codification of the ideas which make up the main structure of psychological theory. This up-to-date basic principles text gives the student a clear picture of today's scientific knowledge of the mechanisms of human behavior.

Kaufmann :
INTRODUCTION TO THE STUDY OF HUMAN BEHAVIOR
By Harry Kaufmann, Hunter College of the City University of New York.
About 200 pages. £1 1s. 0d.
Published March, 1968. Paper.

Designed for collateral reading, this brief new book examines the rationale for, and the methods used in, the scientific study of human behavior.

Millon :
THEORIES OF PSYCHOPATHOLOGY
Essays and Critiques.
Edited by Theodore Millon, Lehigh Univ.
444 pages. Hard Cover £3 5s. 0s.
Soft Cover, £1 16s. 0d.
Published February, 1967.

Rosen and Gregory :
ABNORMAL PSYCHOLOGY
By the late Ephraim Rosen, formerly of the University of Minnesota ; and Ian Gregory, Ohio State University.
553 pages. Illustrated. £3 10s. 0d.
Published February, 1965

Written by a psychologist and a psychiatrist, this unique book integrates psychoanalytic theory, experimental psychology and practical clinical experience—with research in the various biological—and social sciences.

Smith :
DELAYED SENSORY FEEDBACK AND BEHAVIOR
By Karl U. Smith, University of Wisconsin.
109 pages. Illustrated. £2 11s. 0d.
Published September, 1962.

Sherif, Sherif and Nebergall :
ATTITUDE AND ATTITUDE CHANGE
By Carolyn W. Sherif and Muzafer Sherif, Pennsylvania State University ; and Roger E. Nebergall, University of Oklahoma.
264 pages. £3 10s. 0d.
Published March, 1965.

This book delineates the results of the authors' exhaustive study of attitudes—the individuals stand for or against—various persons, groups, ideas, or events. They explain how the individual's attitude defines what he is or is not, what he favors or opposes, and what he accepts or cannot tolerate.

Smith :
MAMMALIAN LEARNING AND BEHAVIOR
By David D. Smith, Bishop's University, Lennoxville, Quebec.
174 pages. Illustrated. £2 15s. 0d.
Published August, 1965.

Here is a revolutionary, yet substantiated, postulation of neurologic events occuring in learning and behavior.

Smith and Smith :
PERCEPTION AND MOTION
By Karl U. Smith, University of Wisconsin; and William M. Smith, Dartmouth College.
341 pages. Illustrated. £3 8s. 0d.
Published October, 1962.

Verville :
BEHAVIOR PROBLEMS OF CHILDREN
By Elinor Verville, formerly of the Central Wyoming Counseling Centre.
567 pages. £3 6s. 0d.
Published March, 1967.

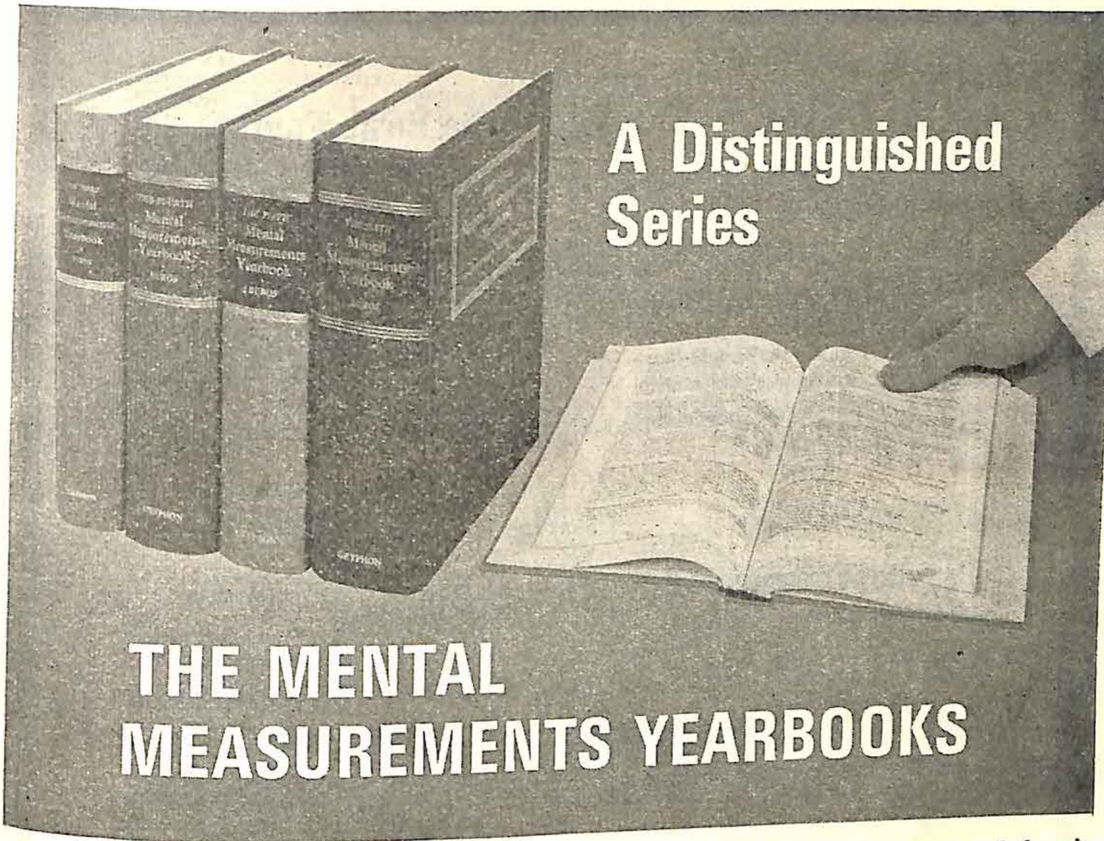
This new text for courses in clinical child psychology will help the student understand behavior problems and prepare him to take an active part in their prevention and treatment.

Whittaker :
INTRODUCTION TO PSYCHOLOGY
By James O. Whittaker, North Dakota State University.
631 pages. Illustrated. £3 4s. 0d.
Published May, 1965.

Here is a superb new psychology text for beginning college students—carefully balanced in content and emphasis to offer an optimal amount of material for the student. The keystone of the book is its strong coverage of motivation, learning and perception.

Also by Whittaker :
STUDENT'S WORKBOOK to accompany INTRODUCTION TO PSYCHOLOGY
224 pages. Paperbound. £1 7s. 6d.
Published May, 1965.

W. B. SAUNDERS COMPANY LIMITED, 12 Dyott St., London, W.C.1



The most widely acclaimed reference works in education and psychology are the encyclopedic *Mental Measurements Yearbooks* with their exhaustive coverage of testing in the English-speaking world. The first five volumes, covering a span of 25 years, have been described as monumental, indispensable and definitive.

The Sixth Mental Measurements Yearbook, published in November, 1965, is by far the largest and best in the series. Reviewers in the professional journals report that all the superlatives used to describe the first five volumes apply even more so to *The Sixth Yearbook*. For example, the reviewer in *Science* wrote: "An evaluation of this last volume can be couched, with palpable redundancy, in terms applied by the reviewers of previous volumes in the series: This remarkable, monumental, colossal, and elaborate book, representing gargantuan and heroic efforts applied compulsively and with missionary zeal is certain to become a 'must' . . . perhaps even the indispensable 'bible' . . . for the increasing number of test users."

Each yearbook . . . issued at intervals of three to six years . . . supplements rather than supplants earlier volumes in the series. Yearbooks still available are the Third (1949, 1062 pages, $18.00), Fourth (1953, 1188 pages, $20.00), Fifth (1959, 1321 pages, $22.50) and *The Sixth Yearbook* (1965, 1752 pages, $32.50) plus *Tests in Print* (1961, 508 pages, $7.00) which serves as a master index to the contents of the first five MMY's.

Purchasers mentioning this advertisement may obtain the 4 MMY's and *Tests in Print* at a special price of $90.

Order directly from **THE GRYPHON PRESS**

220 Montgomery Street Highland Park, New Jersey 08904

Sexual Discord in Marriage

A Field for Brief Psychotherapy

MICHAEL COURTENAY

This book describes and evaluates the technique of focal therapy developed by doctors working in Family Planning Association clinics for marital problems. *Mind and Medicine Monographs No* 16. 35s net

Psychopharmacology

Dimensions and Perspectives

Edited by C. R. B. JOYCE

An international group of specialists present, from the resources of their own disciplines, a review of research methods, clinical and laboratory findings, and sources of reference, contributing to the expanding field of psychopharmacology. *Mind and Medicine Monographs No* 17. 63s net

The Eye

Phenomenology and Psychology of Function and Disorder

J. M. HEATON

This study represents the first systematic attempt to consider disorders of the eye and of vision within a phenomenological and relational framework. *Studies in Existentialism and Phenomenology*. 70s net

The Basic Fault

Therapeutic Aspects of Regression

MICHAEL BALINT

A brilliant contribution to psychoanalytic theory, which develops the concept of the ' basic fault ' as a key to the understanding and interpretation of regression in the analytic situation. 38s net

Social Science Paperbacks

Family and Social Network

Roles, Norms, and External Relationships in Ordinary Urban Families

ELIZABETH BOTT

' Dr. Bott's book represents exploratory research at its best. Systematic definition of key concepts and rigorous adherence to them in analysing the data are achieved even though the entire framework is so new as to be emerging in the course of the study itself .' *Robert N. Rapoport, Man*

No. 35 SSP 17s 6d net ; hardback 35s net

An Outline of Psychology

WILLIAM McDOUGALL

This classic in the field of psychology is now issued for the first time in paperback, with a Foreword by Professor C. A. Mace. *July*, 21s net

TAVISTOCK PUBLICATIONS

SOCIAL SCIENCE
PAPERBACKS

A SYSTEM OF COMPUTER PROGRAMS

for use in Psychology and Education

'A SYSTEM OF COMPUTER PROGRAMS' (by Hallworth and Brebner, price 10/-) is usable even by people who have no experience with computers. For the purpose of analysing data using means and standard deviations and correlations; or a principal components analysis, with Varimax or Promax rotations (and second order factors) and factor scores; or Z-scores, partial correlations, multiple correlations or regressions, this system provides a set of programs which can be used on a computer which uses ICI C-Autocode, ICI K-Autocode, Mercury Autocode or Fortran

The handbook tells you exactly how to prepare your data for the computer and how to interpret the figures that come from the computer

The HANDBOOK is available from the British Psychological Society, Tavistock House South, Tavistock Square, London, W.C.1 (please enclose 10/- with order)

PROGRAM TAPES, CARD DECKS & LISTINGS are available at cost from System of Computer Programs, Education Department, University of Birmingham

COMPARATIVE STUDIES AND EDUCATIONAL DECISION

EDMUND J. KING

Comparative education is an exploding subject, and for the first time Dr. King offers an approach parallel to the latest developments in the social sciences and other comparative disciplines. In doing so he bridges the widening gap between the purely academic world and the world of decision for development. 40s

MODERN TEACHING SERIES

TELEVISION IN EDUCATION

RODERICK MACLEAN

A balanced account of the developing situation as regards the use of television in teaching; the author presents television not as a revolutionary approach to teaching but as a facility through which the teacher can achieve more readily some of those aims that he has traditionally cherished.

13 halftone illustrations. 25s Education Paperback 15s

READINGS IN EDUCATION

W. KENNETH RICHMOND

Including some of the worst as well as the best that has been thought and said on the subject, W. Kenneth Richmond's readings in education have been chosen and arranged so as to evoke an active response from the reader and to provide the material for sustained discussion. September 12th 50s Education Paperback 25s

THE SLOW LEARNER: Some Educational Principles and Policies

Second Edition

M. F. CLEUGH

This new edition has been up-dated to take account of the Mental Health Act of 1959, and a new chapter: 'Seriously Handicapped Children' has been added.

July 11th 25s Education Paperback 12s 6d

METHUEN

THE BRITISH JOURNAL OF PSYCHOLOGY

VOLUME 59, PART 2, MAY 1968

OBITUARY NOTICE: GORDON WILLARD ALLPORT (1897-1967).

R. P. POWER. Simulation of stable and neurotic personalities by subjects warned of the presence of lie scales in inventories.

R. H. THOULESS. Apparent size and distance in vision through a magnifying system.

D. J. MURRAY and BEVERLEY ROBERTS. Visual and auditory presentation, presentation rate, and short-term memory in children.

J. HIGGINS, JUDITH C. PETERSON and LISE-LOTTE LEHD DOLBY. Variability in cognitive control.

M. T. TURVEY. Analysis of augmented recall in short-term memory following a shift in connotation.

A. D. BADDELEY. Closure and response bias in short-term memory for form.

ANNE E. ELLISON and W. E. LAMBERT. Reduction of response interference through verbal repetition.

HANS KREITLER and SHULAMITH KREITLER. Unhappy memories of ' the happy past ' : studies in cognitive dissonance.

J. T. REASON and A. J. BENSON. Individual differences in the reported persistence of visual and labyrinthine after-sensations, and of exponentially decaying visual and auditory signals.

J. SHOTTER. A note on a machine that ' learns ' rules.

30*s. net* ($5.50). *Annual subscription* 90*s. net* ($16.50) *for four issues.*

CAMBRIDGE UNIVERSITY PRESS

The British Journal of Medical Psychology

Volume 41, Part 2. June 1968

Obituary: David M. Kissen

L. Chertok From suggestion to metapsychology

B. W. Murphy Creation and destruction: Notes on Dylan Thomas

Yvonne Blake Psychotherapy with the more disturbed patient

Endre Ugelstadt and C. Astrup A study of psychotic patients treated with individual psychotherapy prior to admission in a mental hospital

H. Halstead Group psychotherapy for the middle-aged

M. Whittet Medico-legal considerations of the A9 murder

H. Kreitler and S. Kreitler Validation of psychodramatic behaviour against behaviour in life

D. J. Smail The place of conscious experience in clinical and medical psychology

E. C. Grant An ethological description of non-verbal behaviour during interviews

T. M. Caine and D. J. Smail Attitudes of psychiatric nurses to their role in treatment

30s. net ($5.50). Annual subscription 90s. net ($16.50) for four issues

Orders should be sent to a bookseller or to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1, England, or Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y.10022, U.S.A.

The British Journal of Social and Clinical Psychology

Volume 7, Part 2. June 1968

P. KING Alienation and the individual
L. J. TAYLOR Alienation, anomie and delinquency
N. F. LEMON A model of the extremity, confidence and salience of an opinion
A. E. PHILIP AND J. W. MCCULLOCH Personal construct theory and social work practice
R. R. HUGGAN A critique and possible reinterpretation of the observed low neuroticism scores of male patients with lung cancer
J. L. HORN AND P. STEWART On the accuracy of clinical judgments
A. W. LANDFIELD The extremity rating revisited within the context of personal construct theory
P. R. MAYO Self-disclosure and neurosis
J. G. INGHAM Comment on 'Effect of training on the performance of the WISC block design subtest' by Josef Schubert
J. SCHUBERT Reply to J. G. Ingham
BOOK REVIEWS

30*s. net* (*USA* $5.50). *Annual subscription* 90*s.* ($16.50)

CAMBRIDGE UNIVERSITY PRESS

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Vol. 20 *No.* 1

April, 1968

ARTICLES

J. K. DUA and R. A. CHAMPION. Contiguity and Reinforcement in Classical Fear Conditioning.
R. G. SALAS and J. F. RICHARDSON. Some Australian Data on Forms A and B of the EPI.
J. F. RICHARDSON. Correlations Between the Extraversion and Neuroticism Scales of the EPI.
D. D. DIESPECKER. An Evaluation of Vibrotactile Signalling Apparatus.
R. W. HODGE and D. D. DIESPECKER. A Two-Channel Signalling Unit for Vibrotactile Experiments.
DAVID T. SIDDLE and GORDON L. MANGAN. Behaviour at the Point of Maximum Approach-Avoidance Conflict.
I. D. JOHN. A Study of Bilateral and Unilateral Spatial Discrimination on the Skin.
E. E. RUMS. Relative Preference as a Function of the Number of Elements in Abstract Designs.
P. J. LIVESEY. A Discrimination Task as a Test of Cognitive Ability in Animals.
R. G. SALAS. Fakeability of Responses on the Eysenck Personality Inventory.
R. M. MOWBRAY and B. M. DAVIES. Scores Obtained by Australian Medical Students on the Eysenck Personality Inventory.
G. N. SEAGRIM. A Binocular Illusion as a Teaching Aid.
GORDON STANLEY. Stimulus—Range and Magnitude Estimates of Distance Based on Object Size.

(*Published by the Australian Psychological Society, University of Sydney, Sydney, N.S.W.*)

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
J. D. NISBET

Assistant Editor :
R. D. DUTCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

L. B. BIRCH
SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. 38, Part 3

Published by
METHUEN AND CO., LTD., 11, NEW FETTER LANE, LONDON, E.C.4,
for the British Journal of Educational Psychology, Ltd.
Registered Offices : 8-24, Sampson Road North, Birmingham, 11. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 8-24, SAMPSON ROAD NORTH, BIRMINGHAM, 11.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor W. SLUCKIN, Department of Psychology, The University, Leicester.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Professor J. D. NISBET, King's College, Aberdeen.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. RALPH HETHERINGTON (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor IAN M. L. HUNTER, Department of Psychology, University of Keele, Keele, Staffs., whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

16 DEC 1968

CONTENTS

G. A. CORTIS: Predicting Student Performance in Colleges of Education . . . 115

N. J. ENTWISTLE and S. CUNNINGHAM: Neuroticism and School Attainment—a Linear Relationship? . . . 123

G. P. GINSBURG and R. G. WHITTEMORE: Creativity and Verbal Ability: A Direct Examination of Their Relationship . . . 133

S. W. HOCKEY: Intelligence, Common Entrance and the General Certificate of Education . . . 140

C. J. PHILLIPS and W. J. BANNON: The Stanford-Binet Form L-M, 3rd Revision: A Local English Study of Norms, Concurrent Validity and Social Differences . . . 148

A. G. DAVEY: The Tristan da Cunhán Children's Concept of Equivalence . . . 162

F. A. HADDON and H. LYTTON: Teaching Approach and the Development of Divergent Thinking Abilities in Primary Schools . . . 171

N. J. ENTWISTLE: Academic Motivation and School Attainment . . . 181

F. B. MURRAY: Operational Conservation of Illusion-Distorted Length . . . 189

RESEARCH NOTES:

D. G. BAGG: The Correlation of G.C.E. A-level Grades with University Examinations in Chemical Engineering . . . 194

A. J. CROPLEY: A Note on the Wallach-Kogan Tests of Creativity . . . 197

J. McLEISH: The Validity and Reliability of Bentley's Measures of Musical Ability . . . 201

P. PUMFREY: The Growth of the Schema of Proportionality . . . 202

D. W. McKERRACHER and R. A. WATSON: Validation of a Short Form WISC with Clinic Children . . . 205

S. H. IRVINE: A Five-year Follow-up of Secondary School Selection Procedures in Central Africa, 1962-67 . . . 208

F. W. SLEE: The Feminine Image Factor in Girls' Attitudes to School Subjects . . . 212

BOOK REVIEWS . . . 215

PUBLICATIONS RECEIVED . . . 230

VOLUME 38 NOVEMBER 1968 PART 3

THE MEASUREMENT OF PARENTAL ATTITUDES AND THEIR RELATION TO BOYS' BEHAVIOUR

BY H. B. GIBSON
(Institute of Criminology, University of Cambridge)

SUMMARY. Separate attitude inventories for fathers and mothers relating to the upbringing of children, were used in a survey concerning 411 primary school boys representative of an urban, working-class area. Analysis of responses produced scales of authoritarianism for both fathers and mothers, and a scale of under-concern for mothers. Scores on these scales were found to be slightly related to bad behaviour in school and later delinquency on the part of the boys. The rate of parental response was rather poor, and it was shown to be significantly related to social handicap and educational and occupational status. Absence of completed inventories was found to be as much related to bad conduct and delinquency as were the attitude scales. The investigation explores the practicability of studying parental attitudes and relating them to objective criteria concerning children.

I.—PREVIOUS WORK.

AN early review of studies relating to parent-child relationships was published by Symonds (1939). This covered experimental and observational work over a fairly wide span of years and represented the general state of theory at that time. The two major dimensions of acceptance-rejection and of dominance-submissiveness had emerged from the studies covered and certain general conclusions were drawn. Children who were ' accepted ' appeared to show characteristics which are generally desirable, whereas ' rejected ' children were reported to be more neurotic, rebellious and delinquent. Dominating parents appeared to have children who were not only more submissive, orderly and polite but better socialized and sensitive also, although more dependent ; by contrast, children of submissive parents appeared to be more aggressive, stubborn and disorderly, although more self-confident. There was also some evidence that child-rearing attitudes had been passed on from the grandparents; dominating parents tended to have been dominated themselves in childhood.

It is within this framework that later studies have been carried on and there is almost a folk lore in the field of parent-child relations which persists to this day. Modifications of the earlier ideas have been suggested by later studies ; first, the ideas concerning authoritarianism in parents have been subjected to much criticism. Read (1945), using an attitude questionnaire designed by Stogdill (1933), found that it was the children whose parents were more libertarian who were the better adjusted. Other workers in this field (Radke 1946, Baldwin 1948, Gough *et al.* 1950, Rosenthal *et al.* 1959) also obtained results which implied that the children of authoritarian parents compared unfavourably with those whose parents were more permissive and democratic. This trend has perhaps reflected the work which has been done on the authoritarian personality by the Adorno school.

Concerning the other dimension of parental attitude, acceptance-rejection, later attention became focussed on what came to be known as ' overprotection ' the attitude of the mother who smothers her child emotionally and perpetuates his infantile, dependent position. Writers such as Levy (1943), Itkin (1952) and Martin (1943), have suggested that overprotection was rather worse for personality development than outright rejection, but of course they were drawing attention to a pathological extreme of attitude.

Two facts about the studies which have been mentioned should be noted. First, most studies have been concerned with abnormal parents and abnormal children, and are not therefore a fair guide to the nature and consequences of the normal variations in parent-child relationships. Second, most of them are based on clinical impressions rather than any sort of objective measures. The studies quoted have included the following methods: direct observation of behaviour, case study methods, interviews of parents and children, questionnaires, attitude inventories, projective tests and rating scales. In these the kinds of questions which are asked and the way in which the clinician interprets the answers depend on his theories about child development, and hence what has been referred to as the 'folk lore' of parent-child relations tends to be self-validating. While even the most objective of techniques depends for what is put into its construction on the theoretical assumptions and hypotheses of the experimenters, later development in this field has been towards more objective forms of measurement. An important landmark is the study of Shoben (1949), who constructed an 85-item attitude inventory by the technique of submitting a pool of 148 pilot items to 50 mothers of problem children and 50 mothers of normal children, and testing which items discriminated significantly between the two groups. Shoben's study did not in fact produce scales by any form of cluster analysis of the items; he relied on the choices of five judges to define his scales. He considered the question of factor analysis of his results but rejected it on the grounds of the great statistical labour that would have been involved.

The advent of easy access to computer facilities has meant that later studies of parental attitudes have been greatly improved by the general technique of intercorrelating all items and defining scales in terms of the clusters which have been indicated by factor or principal component analyses, e.g., Shapiro (1952), Oppenheim (1956), Sears *et al.* (1957), Schaefer and Bell (1958), Zuckerman *et al.* (1958), Schaefer (1961), Salmon (1967). The results of such studies, whether the material was gathered by interviews or by attitude inventories, have shown that parental attitude may be justly represented by the two major dimensions of authoritarianism-permissiveness and overconcern–underconcern, whatever minor variants of these labels have been employed by different workers. The method of cluster analysis applied to attitude inventories undoubtedly ensures construct validity for the attitudes which are measured. The problem arises in establishing the probability that certain specified types and degrees of attitude towards child-rearing practices in parents will be associated with certain behavioural characteristics in their *own* children. Where the populations studied have been abnormal, as with the parents of children attending child guidance clinics, and the attitudes extreme, valid associations have generally been demonstrated, but the demonstration of such associations in normal populations is another matter. It is incorrect to generalize about the relations between parental attitudes and the behaviour of children when one is testing only *selected* populations, such as those who seek help at child guidance clinics. Very different relationships may exist in normally occurring populations.

Although the validity of published scales of parental attitude has been questioned by quite a number of studies in recent years (Gordon 1957, Leyton 1958, Medinnus 1963) there is little indication in published research that efforts have been directed into pursuing the enquiry among normally occurring populations.

II.—THE PRESENT STUDY.

The present study concerns 411 boys first contacted at the age of 8 years in primary schools, as representative of an urban working-class area with a high rate of delinquency and other social problems. These boys are the subject of a longitudinal psycho-social survey. It was desired to measure parental attitudes by means of inventories, and to investigate the degree to which such attitudes might be associated with behaviour problems in the boys and the later development of delinquency.

Statements used by Shoben (1949) and Oppenheim (1956) were considered with reference to the two postulated dimensions of authoritarianism-permissiveness and overconcern-underconcern. Modifications of wording were made in the direction of brevity and of comprehensibility by working-class parents of minimal education.

Considerable pilot work was done with four provisional forms of attitude inventories, involving approximately 400 working-class parents who were not members of the experimental population. As a result of this work separate final forms of inventory, each of 33 items, were prepared for mothers and fathers.

The final inventories were sent by post to the majority of the parents of the 411 boys in the experimental population, with a letter requesting both the mother and father to complete their own form and return them in stamped envelopes provided. In a minority of cases, the forms were not sent but presented by family interviewers, who thought that in certain cases a personal explanation might be more productive of co-operation. In the case of 57 boys, no maternal attitude schedule was given, and for 87 boys no paternal attitude schedule was given. The reason for this non-presentation of the form included previous death or desertion of the parent, and known illiteracy or foreign speech, but in many cases the interviewers were unable to establish sufficiently good contact with the families even to ask them to fill in such inventories without provoking hostility. A substantial minority of the parents in this study co-operated only rather grudgingly, and were somewhat averse to filling in forms of any kind.

III.—RESULTS.

Because of the extremely slow rate of return of the forms, analysis was undertaken when 188 schedules only of maternal attitude, and 173 schedules only of paternal attitude had been received. Both schedules were analysed by the method of principal components.

Mothers' inventory.

The first two components extracted accounted for 8·4 and 6·8 per cent of the total variance. The first component was identified as one of authoritarianism–permissiveness and the second overconcern–underconcern, by the nature of the wording of the items. Items which had loadings of $< \cdot 2$ on either component were disregarded. The resulting scales were 12 items having a mean loading of ·42 on the first component, and 15 items having a mean loading of ·33 on the second component.

Fathers' inventory.

The first two components extracted accounted for 10·4 and 6·5 per cent. of the total variance. The major component was identified as being concerned with authoritarianism, by the wording of the items with high loadings on it, e.g., "Strict discipline develops a good strong character in children." The second component appeared to be concerned with a 'reasonable' social attitude and most of the buffer items had high loadings on it, such as the buffer item, "Immunization against polio should be compulsory." Thirteen items (including those which were deliberately phrased as buffers) had loadings of $< \cdot 2$ on the main component and were disregarded. The resulting authoritarianism–permissiveness scale of 20 items had a mean loading of ·38 on the main component.

Relations between the scales.

There was no significant association between underconcern on the maternal scale, and authoritarianism on either the maternal or paternal scales. There was, however, a significant association between the two authoritarian scales completed by the mother and the father ($C = \cdot 35$; $p < \cdot 001$).

Relation of the attitude scales to variables concerning the parents.

All completed schedules were scored in terms of the scales which had been produced by the principal component analysis. There was considerable delay in obtaining completed schedules from many mothers and fathers, and even after repeated reminders it was impossible to get a good rate of return. Of the 411 boys in the sample, completed inventories were eventually obtained from 291 of their mothers and 232 of their fathers. In considering this rather low rate of return it must be remembered that, as mentioned before, in 57 cases of maternal and 78 cases of paternal forms, either there was no parent to ask or the parent could not be asked for various reasons. It is important to consider our population as 411 however, as that figure represents the boy population representative of the area studied, and this research shows how far it is *practicable* to study parental attitude by this means and relate it to other variables. For convenience, the categories of 'no parent,' 'not asked' and 'asked in vain' have been thrown together in Table 1 to illustrate just what are the chances of getting completed attitude schedules from the parents of boys in a population such as this.

TABLE 1

ASSOCIATIONS BETWEEN PARENTAL ATTITUDES AND TEACHERS' RATINGS OF THE BOYS' BEHAVIOUR.

Parental Attitude	Total	Behaviour Problem Boys	
Mother Authoritarian :	n	n	%
Low	59	3	5·1
Average	178	27	15·2
High	54	10	18·6
No form	120	24	20·0
Mother Underconcerned :	n	n	%
Low	48	5	10·4
Average	200	25	12·5
High	43	10	23·3
No form	120	24	20·0
Father Authoritarian :	n	n	%
Low	31	2	6·5
Average	144	19	13·2
High	57	8	14·0
No form	179	35	19·6
N	411	64	

There was a rather better chance of getting completed inventories from the boys' mothers (70·8 per cent) than from their fathers (56·4 per cent), and this rate of completion was closely related to the status of the *mother* of the family. Both parents were rated on a variable of combined education and social class, and it was found that the rate of return of completed inventories, both maternal and paternal, was related to the mother's level ($p < \cdot 001$) rather than the father's. There was no overall significant tendency for authoritarianism in mothers or fathers to be related to their level of education and occupational status (as has been suggested by some previous studies). There was, however, a non-significant trend for the mothers of the highest status group to be relatively less authoritarian.

Families were divided into three levels of social handicap on objective criteria of economic standards, housing, family size, etc. (Details of this and other social variables are given by West, 1968). The chance of getting inventories completed by fathers was considerably reduced by the degree of social handicap present in the family: 65·9 per cent of fathers completed an inventory, where there was no social handicap, but where there was a ' severe ' handicap, only 36·4 per cent of fathers responded ($p < \cdot 001$). The lowered rate of return of maternal inventories associated with social handicap, was not statistically significant. Maternal overconcern tended to be more characteristic of those mothers with no social handicaps, but the relationship was not significant. In mothers a highly authoritarian attitude was slightly related to severity of social handicap: in fathers an authoritarian attitude had no such relationship.

Parental attitudes as related to the boys' characteristics.

All the boys were rated on behaviour in school by their school teachers at the ages of 8+ and 10+ on a standard form which has been described elsewhere (Gibson 1964). On this form, boys are rated A, B or C by summation of weighted scores. Thus boys who have been rated C by the judgments of two school teachers at the separate age levels can be characterized as behaviour problems; 64 boys were so rated. The relationship between behaviour problems and the attitudes of the parents, or the fact of the parents being unassessable by means of the attitude inventories, is shown in Table 1. Here it may be seen that there is a tendency for behaviour problems to be associated with high authoritarianism and a high degree of underconcern in mothers, and with high authoritarianism in fathers. The fact of there being no form returned is also recorded in this table, and it may be seen that this accords with a bad prognosis for behaviour problems. The values of χ^2 do not quite reach the ·05 level, but the percentages indicate a fairly linear trend in the expected direction for each of the three scales.

When both parents completed the inventory (N=224) only 12·5 per cent of cases were behaviour problems, as compared with 20·0 and 19·6 per cent behaviour problems associated with non-return by mothers and fathers respectively, as shown in the table.

A further piece of objective evidence about the boys, which was available some years after the attitude inventories had been completed, was whether they had or had not appeared before the juvenile court on any charge up to the age of 13 years. A search of the relevant records revealed that 25 of the boys had in fact appeared before the courts on at least one occasion. Of these 25 cases, attitude inventories had been completed by both parents in only eight (32·0 per cent) cases as compared with 56·0 per cent in the non-delinquent part of the sample ($p < \cdot 05$).

Mothers of 14 boys who appeared before the courts had completed inventories, and even with this small number there is a significant association with authoritarian attitude when the scale is dichotomized ($p < \cdot 02$). There is a positive association (N.S.) between maternal under-concern and court appearance of the boys. Only 10 of the delinquent boys had fathers who completed the inventory, and of these cases nine scored above the median of the authoritarian scale.

IV.—DISCUSSION.

This study set out to measure the attitudes towards child rearing in a population of urban, working-class parents and to relate such attitudes to independent measures of their sons' behaviour. The study aimed to explore the practicability of this venture. Quite a high percentage of the parents did not complete the inventories for one reason or another, and this non-completion was as closely related to bad behaviour and delinquency in the boys as any attitudinal measures that could be obtained from the inventories. Zuckerman (1960) found in a small study that non-return of attitude inventories by parents related significantly to a scale of the maladjustment of their children, and this result parallels the present finding. In the present case it may be argued that had the interviewers been more bold, persistent or skilled, a higher rate of completion of inventories would have been obtained; the parental characteristics which intimidated the interviewers are obviously such as are associated with (and perhaps causative of) bad behaviour in school and later delinquency on the part of the boys. This fact may be a valid argument against the attempt to measure attitudes by means of inventories when parents are unco-operative and hostile, but it is an equally valid argument aginst the attempt to measure attitudes by interview when such parents avoid giving information to interviewers.

It was not surprising to find that the return of inventories was significantly related to the variables of social handicap, education and occupational status of the parents. What was less expected was the importance of the education and status of the mother of the family rather than the father. It may be that the completion and return of inventories depended more upon the mothers' concern and initiative than upon the fathers', as the matter concerned child-rearing practices. It was also somewhat unexpected to find so little relationship between the scales of attitude and the variables of social handicap, education and status.

Measuring underconcern on the part of mothers by means of such an inventory may indeed be a hopeless task, for if the mother simply cannot be bothered with her children, she is likely to express her attitude of underconcern by refusing to fill in the inventory. While these results have shown a moderate association between uuderconcern and both maternal and paternal authoritarianism on the one hand, and bad behaviour and delinquency in schoolboys on the other, even these findings must be considered with reservation. The reason for such reservation lies in the high rate of non-return among certain categories of parents. Thus, as completed inventories were available concerning only 49·1 per cent of boys who had fathers in the lowest category of educational and occupational status, we know nothing about the relationship between boy's behaviour and father's attitude for the other half of the sample in that grade.

The results of this study have tended to confirm the evidence from earlier studies that an authoritarian attitude and underconcern on the part of parents are slightly related to behaviour problems in the children; this is not a contribution of any great intrinsic value. This study has not thrown any light on the question of maternal overconcern, but this is understandable because the two characteristics of the boys considered were bad behaviour of an acting-out variety. In demonstrating associations, no causative hypothesis has been advanced. The value of this study lies in its methodological contribution to the area of psycho-social research into parent-child relationships, and exploration of what is practicable, given a population representative of a socially deprived area.

V.—REFERENCES.

BALDWIN, A. L. (1948). Socialization and parent-child relationships. *Child Development*, 19, 127-136.

FITZ-SIMONS, M. J. (1935). *Some Parent-child Relationships as shown in Clinical Case Studies.* New York: Teach. Coll., Columbia Univ.

GIBSON, H. B. (1964). The Spiral Maze; a psychomotor test with implications for the study of delinquency. *Brit. J. Psychol.*, 54, 219-255.

GORDON, J. E. (1957). The validity of Shoben's Parent Attitude Survey. *J. Clin. Psychol.*, 13, 154-156.

GOUGH, H. G., HARRIS, D., and MARTIN, W. (1950). Children's ethnic attitudes II: relation to parental beliefs concerning child training. *Child Development*, 21, 169-181.

ITKIN, N. W. (1952). Some relationships between intra-family attitudes, and pre-parental attitudes towards children. *J. Genet. Psychol.*, 80, 221-252.

LEVY, D. M. (1943). *Maternal Overprotection.* New York: Columbia Univ. Press.

LEYTON, D. A. (1958). A study of the validity of parent attitude measurement. *Child Development*, 29, 515-520.

MARTIN, A. R. (1943). A study of parental attitudes and their influence upon personality development. *Educ. Boston*, 63, 596-608. (Cited by Itkin.)

MEDINNUS, G. R. (1963). The relation between parent attitudes and parental acceptance of the child. *J. Genet. Psychol.*, 103, 117-121.

OPPENHEIM, A. N. (1956). *Inventories for Child Rearing Attitudes and Parental Behaviour.* Unpublished study.

RADKE, M. (1946). The relation of parental authority to children's behaviour and attitudes. *Inst. Child Welfare Monogr.*, 22, Univ. Minnesota Press.

READ, K. H. (1945). Parents' expressed attitudes and children's behaviour. *J. Consult. Psychol.*, 9, 95-100.

ROSENTHAL, M. J., FINKELSTEIN, M., and ROBERTSON, R. E. (1959). A study of mother-child relationships in the emotional disorders of children. *Genet. Psychol. Monogr.*, 60, 65-116.

SALMON, P. (1967). *A Study of the Social Values and Differential Conformity of Primary Schoolboys as a Function of Maternal Attitude.* Ph.D. Thesis, Univ. of London.

SEARS, R. R., MACCOBY, E., and LEVIN, H. (1957). *Patterns of Child Rearing.* New York: Row, Peterson.

SCHAEFER, E. S., and BELL, R. Q. (1958). Development of the Parental Attitude Research Instrument. *Child Development*, 29, 339-361.

SCHAEFER, E. S. (1961). Converging conceptual models for maternal behaviour and for child behaviour. In Glidewell, J. C. (Ed.), *Parental Attitudes and Child Behaviour.* Springfield: Thomas.

SHAPIRO, M. B. (1952). Some correlates of opinions on the upbringing of children. *Brit. J. Psychol.*, 43, 141-149.

SHOBEN, E. J. (1949). The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 39, 101-148.

STOGDILL, R. M. (1933). Attitude of parents, students and mental hygienists towards children's behaviour. *J. Soc. Psychol.*, 4, 486-489.

SYMONDS, P. M. (1939). *The Psychology of Parent-child Relatinoships.* New York: Appleton-Century.

WEST, D. J. (1968). *Behaviour Problems in Primary Schoolboys.* London: Heinemann (in the press).

ZUCKERMAN, M., RIBBACK, B. B., MONASHKIN, I., and NORTON, J. A. (1958). Normative data and factor analysis on the Parental Attitude Research Instrument. *J. Con. Psychol.*, 22, 165-171.

ZUCKERMAN, M., BARRETT, B. H., and BRAGIEL, R. M. (1960). The parental attitudes of parents of child guidance cases. *Child Development*, 31, 401-417.

(*Manuscript received* 23*rd October*, 1967)

A STUDY OF SOME DIFFERENCES BETWEEN BACKWARD AND NORMAL READERS OF AVERAGE INTELLIGENCE

By K. LOVELL and A. GORTON
(*University of Leeds Department of Education*)

SUMMARY. A group of 50 backward readers aged 9 to 10 years were selected from a total age group of 697 pupils in 16 primary schools of a small local education authority. A battery of tests was given to this group and to a control group of 50 pupils of normal reading ability. The test scores were intercorrelated, separately for the two groups, and in each case a Principal Components Analysis was carried out with rotation to a varimax criterion. The factor patterns of the two groups were different. In the backward readers, but not in the normal readers, reading age was clearly linked with test performance, which seems to depend upon the integrity of central processes. Moreover, the study throws some light on the previous literature describing the results of single tests given to backward readers and matched groups of normal readers, which tends to be inconclusive.

I.—INTRODUCTION.

THERE is evidence in the literature dealing with reading failure that in some instances the backwardness is due to a specific impairment or lack of integrity in the central nervous system. Difficulty in reading may in these cases be but one sympton of a wider syndrome with writing often being affected as well. On the other hand there is no wish to minimise the effects of poor and inconsistent teaching, poor cultural background, or emotional maladjustment on reading failure. These influences may well contribute to the persistence of backwardness even when they are not the root cause of the failure. Lovell (1966) has outlined elsewhere some of the evidence which led us to believe that reading failure in some children, at least, might be due to a specific failure in central processes. The literature in this field is extensive and cannot be reviewed here. Some teachers and psychologists would support this view, as would many paediatricians and neurologists. Again, it must be stressed that because reading failure may be found to be linked with a specific impairment in the central nervous system it does not mean that remedial teaching is necessarily ineffective. But it could mean that in our present state of knowledge, less progress might be made with these pupils than teachers would wish.

Our hypothesis was that there would be differences between backward and normal readers, of average intelligence, on visuo-spatial and neuro-psychological tests in the direction of indicating a greater degree of neurological impairment in the backward readers.

II.—PLAN OF THE STUDY.

(i) *Subjects*. The investigation was carried out amongst 9–10-year-olds in the schools of a small local education authority. This age range was chosen for two reasons. First, it is now thought that visual perceptual abilities may have become less important to reading progress by 9 years of age, while variations in auditory perceptual abilities and/or the ability to transpose between audition and vision have become more important (Geschwind, 1964). Second, by 9 years of age children have been at school long enough for reading disability to be clearly recognisable. Many children aged 7 and 8 years of age make considerable progress under remedial teaching, but good evidence has been provided by Lovell (*op cit*). and by Morris (1966) that only a minority of those who are backward in reading at 9 succeed later on in reaching the attainment of their age group, and very few indeed become good scholars.

The backward readers were selected as follows :

(*a*) All children in the age group attending the authority's schools were given the Schonell Silent Reading Test B as a group test. The SD of the test calculated from this sample was 1 year 6 months. Using Schonell's reading age norms we selected as possible subjects all those whose scores were such that their reading ages were more than 1 SD below the norm.

(*b*) These possible subjects were then given individually the Vernon Revision of the Burt Graded Word Reading Test, the test also being given to all the children in the classes from which the possibles were drawn. We thus had a representative sample of children in respect of reading achievement, and from this sample we were able to calculate our own mean and SD for the Burt Graded Word Reading Test. Reading ages were calculated in accord with the test instructions, and the group of backward readers selected were those whose reading age was more than 1 SD below that which was normal for their age. All these backward readers were then given the Terman Merrill revision of the Binet Scale, and those with an IQ of less than 85 were eliminated from the sample.

Using these criteria, 131 backward readers were obtained from the 16 schools out of a total of 697 pupils in the age group. From the 131 pupils, 50 were selected at random for detailed study. By these means 7·17 per cent of the age group were selected as backward readers for the study. The sample consisted of 33 boys and 17 girls ; this ratio of two boys to one girl is consistent with results found in other investigations.

A control group of 50 pupils were selected in the following way. From the classes of pupils who had been given Vernon's revision of the Burt Graded Word Reading Test, pupils were considered who were not backward in reading. These were grouped according to age, sex and time spent in the school in which they were located. Those who matched the members of the experimental group in respect of these three variables were given the individual intelligence test. From those scoring an IQ of 85 or more, pupils were selected at random to make up a control group of 50. Basic data for the backward and control groups of readers are given in Table 1.

In the LEA in question pupils had been introduced to reading by a ' look and say ' approach. After two or three terms in the infant school this was gradually supplemented by a phonic approach and all of the backward readers had experienced both approaches.

(ii) *Tests used.*

(1) The Illinois Test of Psycholinguistic Abilties. This was administered and scored according to the procedures of McCarthy and Kirk (1961). For each subject it yields a language age.

(2) Test of auditory-visual integration as administered and scored by Birch and Belmont (1964). The subject had to identify a visual pattern of dots which corresponds to the patterning of a rhythmic auditory stimulus.

(3) Monroe's (1932) test of sound-symbol association administered and scored as by Monroe.

(4) A test of auditory discrimination. Pairs of words were presented orally and the subject had to say if these were of the same sound or diffeeent. The test had been earlier standardized by Gorton (1964).

(5) Visual-motor Gestalt test. The eight cards used were those described by Bender (1938) and the test was administered according to her instructions. The marking scheme adopted allowed 1 mark for a scribble or totally incorrect reproduction; 2 marks for a general resemblance of the form of the model; 3 marks for a clear resemblance of the model form but inaccurate relationships; 4 marks for a clear resemblance of the model form but spatial disorientation; and 5 marks for a good reproduction in respect of size, shape and relationship.

(6) Shapiro's test of Rotation (Shapiro *et al*, 1962). The 'rotation effect' refers to a tendency to reproduce abstract designs correctly but in a wrongly orientated position.

(7) Spatial orientation test designed by Semmes Weinstein *et al.* (1955). The subject is required to follow on the ground, routes represented on maps. Five of the maps were perceived visually; nine others, which were rotated and mirror images of the visual maps, were perceived through touch alone.

(8) Test of left-right discrimination. Fifteen items were used to assess three aspects of left-right discrimination; viz., the execution of uncrossed commands, the execution of crossed commands with eyes open and with eyes closed, and the identification of lateral parts of the person confronting the subject. Thus, the test included instructions of the type: with eyes open indicate the right-hand and use it to touch the right knee; with eyes closed point with the left hand to the right eye; point with the right hand to the left eye of the person facing. This test combines items drawn from the work of Benton (1959), Harris (1957) and Herman (1959).

(9) Motor impairment test. Use was made of Stott's (1966) version of the Göllnitz Revision of the Oseretsky Test of Motor Development. For each subject a score of 'motor impairment' was obtained. The test appears to measure neurological aspects of motor failure while minimizing those due to differences in stature or muscular strength. Moreover, the items seem to be free of sex bias. When giving the tests, rather than rely on verbal instructions alone, the experimenter performed each task in front of the child and then asked him to perform it.

Except for the test of motor impairment all the tests, in either the same or similar forms, had been used by the senior author (K.L.) in published or unpublished studies of reading failure (e.g., Lovell 1964a, 1964b) although they had never been used in one and the same study. Moreover, a strong case can be made out for suggesting that some of the tests do reflect neural impairment. For example, the Shapiro test has been shown (Shapiro *et al.* 1962) to yield differences in score between certain kinds of brain injured patients and normal adults.

In addition to the tests given above both groups were given the Language Modalities test for Aphasia (Wepman and Jones, 1961) and observations were made throughout the testing sessions for choreiform symptoms (Prechtl and Stemmer, 1962).

III.—THE RESULTS.

Table 1 shows the means and standard deviatious in respect of chronological age, IQ, reading age, and the scores on the nine tests for both the backward and normal readers. Significant differences in the means (P= ·05 or less) in favour of the normal readers were found in the case of: reading, age, language age, auditory-visual integration, sound-symbol association, spatial orientation, left-right discrimination, and the test of motor impairment.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR TWLEVE VARIABLES.

Variable	Backward Readers		Normal Readers	
	Mean	SD	Mean	SD
Chronological age (months)	116·1	4·1	114·7	3·3
IQ	96·1	5·0	98·5	4·0
Reading age (months)	82·8	9·6	115·5	3·1
Language age	101·4	9·0	106·9	3·2
Auditory-visual integration	6·9	2·4	8·3	1·2
Sound-symbol association	6·3	2·9	7·9	1·3
Auditory discrimination	17·3	4·0	18·2	1·0
Visual-motor Gestalt	29·3	4·5	30·3	3·3
Shapiro rotation	9·3	5·5	9·1	2·3
Spatial orientation	47·2	11·5	62·5	7·0
Left-right discrimination	9·9	3·1	12·4	1·3
Motor impairment	–3·3	3·8	–2·0	1·5

Differences between the group means were not significant, however, in the case of age, IQ, and the tests of auditory discrimination, visual-motor Gestalt and Shapiro rotation, although in the case of the three tests the differences were in the expected direction. It was obvious, however, that little more could be obtained from a consideration of the means and SDs alone. Accordingly inter-correlations between the twelve variables were calculated within both the backward and normal groups, with motor performance or the reverse of motor impairment being used as one of the variables (Table 2).

TABLE 2

INTERCORRELATIONS OF THE TWELVE VARIABLES: BACKWARD AND NORMAL READERS.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1 Chronological age	.	09	70	04	15	–01	14	–01	–19	07	–19	–06
2 IQ	–12	.	06	–08	–06	–32	17	–01	–12	03	18	17
3 Reading age	36	–04	.	21	06	11	13	09	–15	15	–22	–16
4 Language age	36	–05	48	.	44	56	23	16	14	10	–15	–03
4 Auditory-visual integration	31	–10	56	58	.	25	22	01	03	07	–06	05
6 Sound-symbol association	19	–15	35	50	33	.	–03	09	–16	13	–32	–01
7 Auditory discrimination	24	–09	38	75	43	42	.	–12	08	40	15	–03
8 Visual-motor Gestalt	28	10	23	37	46	23	25	.	–05	–10	19	–31
9 Shapiro rotation	25	–02	57	59	75	16	42	29	.	–03	15	18
10 Spatial orientation	19	05	53	63	71	38	52	46	64	.	–36	–09
11 Left-right discrimination	13	08	37	56	66	34	50	41	64	61	.	01
12 Motor performance	14	–08	55	67	79	44	53	48	76	65	66	.

Coefficients for backward readers below the diagonal; for normal readers above the diagonal. Decimal points omitted.

The reason for calculating the intercorrelation matrices is that test scores from different groups may differ not only in respect of their means and SDs, but also in regard to their intercorrelations. Indeed a difference between the groups was immediately apparent; the mean intercorrelation coefficients between the variables being +39 and +15 respectively for the backward and normal readers.

Table 3 shows the results of a Principal Components Analysis of the two intercorrelation matrices obtained using the University of Leeds KDF9 Computer.

TABLE 3

ANALYSIS OF THE VARIABLES: UNROTATED AND ROTATED AXES.

Variable	Principal Compoenents Axes						Rotated Axes					
Backward Readers	I	II	III	IV	V	VI	I	II	III	IV	V	VI
1 Chronol. age	38	–50	00	–70	11	–17				93	18	16
2 IQ	–05	73	–45	–28	40	–04		99				
3 Reading age	68	–16	16	–14	46	31	68		36	45	–21	
4 Language age	83	–17	–25	09	06	–25	46		27	22	13	71
5 A-V. integr.	86	08	29	–04	–08	10	84	–11	12	15	26	14
6 Sound-symbol	53	–29	–56	17	–02	48	14		92		12	26
7 Aud. discrim.	69	–23	–31	26	05	–43	29		19	11		87
8 V–M Gestalt	54	23	–16	47	–57	05	29		11	17	88	
9 S. rotation	80	17	40	03	17	–09	89			14		22
10 Sp. orient	82	17	–01	03	01	04	70	10	23		23	31
11 L–R. discrim.	77	29	00	14	–10	–12	66	10		–11	29	41
12 Motor perf.	87	12	13	16	–11	13	81		25		24	26
Percent. var.	48	10	8·4	8·1	6·4	5·5	33	8·6	10·3	10·2	9·7	14·8
Normal Readers												
1 Chronol. age	52	64	–13	02	34	21		92				
2 IQ	–16	54	32	22	12	–53			15	72	25	–38
3 Reading age	63	53	–18	08	24	25	12	89			–12	
4 Language age	65	–45	19	28	20	–10	87					
5 A–V. integr.	50	–27	26	29	19	–11	69		12			
6 Sound-symbol	61	–54	–16	–14	09	–26	70			–47		–25
7 Aud. discrim.	32	20	68	29	–31	01	23		75	36		16
8 V–M. Gestalt	10	–11	–48	63	–14	–06	20		–25	19	–73	
9 S. rotation	–20	–32	45	16	16	67					14	90
10 Sp. orient.	44	14	38	–27	–61	01			85	–25		–12
11 L.–R. discrim.	–51	04	13	69	06	–03		–20	–20	74	–22	28
12 Motor perf.	–23	–06	45	–35	60	–19			–19	14	84	
Percent. var.	20	14·5	12·8	11·8	9·5	8·0	15·6	14·5	12·4	12·9	11·6	9·7

Decimal points omitted.

Six components were extracted in the case of both groups; the components accounting for 86 per cent of the variance of the scores in the case of the backward readers, and 77 per cent in the case of normal readers. In addition, the standard error of the saturation coefficients in respect of each of the six factors was calculated using the formula proposed by Burt (1952, p. 128, formula XXVI) and are given below:

	I	II	III	IV	V	VI
Backward readers	·09	·15	·16	·18	·19	·21
Normal readers	·13	·15	·16	·17	·19	·21

The principal Components were rotated to a varimax criterion and these results are also shown in Table 3, with loadings lying between ± ·10 omitted.

In order to compare the factor pattern of the backward readers with that of the normal readers, the method of Amhavarra* was used. In essence we have to calculate the rotating matrix f so that

$$[F_B]\ [f] = [F_N]$$

where F_B and F_N represent the matrix of factor loadings for the backward and normal readers respectively. Table 4 shows the results.

TABLE 4

ROTATING MATRIX TO COMPARE THE FACTOR PATTERNS OF THE BACKWARD AND NORMAL READERS.

		Normal Readers Rotated Factors					
		I	II	III	IV	V	VI
Backward Readers Rotated Factors	I	041	—089	018	175	541	816
	II	—179	—273	269	789	145	—042
	III	—581	100	081	582	008	554
	IV	051	974	077	—103	—136	—114
	V	—164	—384	418	—270	722	237
	VI	—348	—227	—717	—495	—029	—259

Decimal points omitted.

IV.—DISCUSSION.

In the case of backward readers the first factor above accounts for 48 per cent of the variance and suggests a dimension of neurological integrity-impairment which is rather specific in nature. Although it has an almost zero correlation with IQ, it has correlations exceeding ·65 with reading age, and the tests of auditory-visual integration, Shapiro rotation, spatial oreintation, left-right discrimination and motor performance. The high correlation of ·84 between the factor and the test of auditory-visual integration is entirely consonant with the findings of Rudnick, Sterritt and Flax (1967) and also of Birch and Belmont (1964), that performance on the test is associated with reading ability in the case of younger pupils, but in 9 and 10-year-olds only in the case of backward pupils. In our group of normal readers there is a near zero correlation between reading ability and the score on this test.

Factor II in the group of backward readers indicates intelligence as measured by the Terman Merill revision of the Binet Scale. Within the IQ range of 85 to 113 it correlates with little else in this test battery. The third factor indicates an ability underlying performance on the sound-symbol association test, its correlation with reading age being only ·36. In the case of factor IV, which clearly reflects a dimension of age-youth, there is only a moderate sized correlation (·45) with reading age; the relationship being very different in size from that between chronological and reading age in normal readers. Factor V suggests an ability which underpins performance on the visual-motor Gestalt test and it has only small correlations with the other variables. This finding is entirely consonant with the findings of Werner *et al.* (1967) that there are only small correlations between group Bender-Gestalt error scores and both reading grades and reading test scores. The last factor, however, is of interest for while it indicates an ability underlying performance on the test of auditory discrimination, it also correlates ·71 with language age as measured by the Illinois Test of Psycholinguistic Abilities, ·41 with the test of left right discrimination, but not all with reading age.

---

*A programme was kindly made available by Professor Max Hamilton and Mr. R. J. McGuire of the Department of Psychiatry, University of Leeds.

The analysis has explained some 84 per cent of the variance in reading age among the backward readers ; some 46 per cent being accounted for by the first factor of neurological integrity, 20 per cent by age and 13 per cent by the ability to associate sound and symbol.

In the case of the normal readers the position is very different. The first factor can be labelled language ; it has substantial correlations with the tests of auditory visual integration (·69) and sound symbol association (·70) but virtually no correlation with reading age. Factor II suggests an age-youth dimension, and in the case of these normal readers it correlates highly (·89) with reading age. Indeed, this factor accounts for almost all the identified variability in reading age. The remaining factors suggest rather specific abilities. Factor III links performance on the spatial orientation task with that on the test of auditory discrimination, while the fourth factor suggests a linkage between the ability to make left-right discriminations and intelligence in normal children. The last two factors seem to indicate abilities which underpin motor performance, and competence in the Shapiro rotation test, respectively. Thus, in respect of normal readers too, our findings which relate to the visual-motor Gestalt test and reading age are consonant with the findings of Werner *et al.* (*op. cit.*).

Table 4 shows that there is a high correlation only between Factor IV found among the backward, and Factor II found among the normal, readers. The figure of ·974 indicates that the angle between the vectors is approximately 13 degrees and we are clearly justified in labelling each as an age-youth dimension. In the case of the other factors there is no correspondence at all.

It is clear overall that the test battery did little to explain the variation in reading age in the case of normal readers, although it does throw considerable light on this problem in backward pupils.

Calculations were also made of each child's factors scores on the six factors found in the group of backward readers. The five children with the lowest scores on Factor I all showed the spasmodic jerky symptoms described by Prechtl and Stemmer (1962) ; the only other child showing these symptoms had a low score on Factor VI or the factor which has its highest loadings on the tests involving auditory discrimination and language age. No child in the group of normal readers showed any of the symptoms described by Prechtl and Stemmer.

Of the eight children who had the lowest scores on Factor VI, four were classified as aphasic in terms of the Language Modalities Test for Aphasia. No other child could be classified as aphasic on the L.M.T.A. either in the backward or normal group of readers.

Thus 12 per cent of the group of backward readers or 1·7 per cent of the age group showed evidence of Prechtl's symptoms. This in no way implies that the other backward readers did not have some degree of neurological impairment, but it does give confirmation, independent of the factor analysis, of such impairment in these instances. Moreover, it is possible that had pupils with IQs less than 85 been included in the study, the incidence of neurological impairment would have been higher.

Factor I in the group of backward readers is clearly linked with auditory-visual integration and motor performance as well as with performance on the tests of Shapiro rotation and spatial orientation. This finding is entirely consonant with findings of Geschwind (1964) who came to the conclusion on the basis of clinical neurological evidence that learning to read is greatly dependent on auditory-visual connections particularly *after* the earliest stages. If auditory functions become increasingly important in reading 9–10 years of age then deaf children should find it very difficult to read beyond the 9-year-old level. This expectation has been borne out by the survey of Myklebust (1960) in the U.S.A., who found that the profoundly deaf child at 18 is generally below that of an average child of 9 in respect of reading skills.

It will be observed that the word dyslexia has not appeared in this paper up to this point. This was deliberate. Rather it has been shown that when a batch of backward readers is selected from the junior schools of a LEA, statistical evidence strongly suggests that reading age can be classified, using factor analysis, among tests of auditory-visual integration and motor performance. This does not prove that reading failure is caused by some specific neuorological impairment, but it does provide evidence that in many backward children, such impairment and reading skills are linked. The study suggests that it is the greater clustering of the test scores in the backward, but not in the normal, readers which throws some light on the aetiology of reading failure at this age level. It also helps us to undertand why the literature tends to be confused when single tests are given to normal and backward readers and only the differences between the means of the groups considered.

V.—REFERENCES.

Bender, L. (1938). A visual motor Gestalt test and its clinical use. *Res. Monog. Amer. Ortho-psychiat. Assoc.*, No. 3.

Benton, A. L. (1959). *Right-Left Discrimination and Finger Localization.* New York: Hoeber-Harper.

Birch, H. G., and Belmont, L. (1964). Auditory-visual integration in normal and retarded readers. *Amer. J. Orthopsychiat.*, 34, 852-861.

Burt, C. (1952). Tests of significance in factor analysis. *Brit. J. Psychol., Stat. Sec.*, 5, 109-133.

Geschwind, N. (1964). The development of the brain and the evolution of language. *Monogr. Series on Language and Linguistics*, No. 17.

Gorton, A. (1964). *An Investigation of Reading Difficulties in a Sample of Children.* Unpublished M.A. Thesis, University of Reading.

Harris, A. J. (1957). Lateral dominance, directional confusion and reading disability. *J. of Psychol.*, 44, 283-294.

Hermann, K. (1959). *Reading Disability.* Copenhagen: Munksgaard.

Lovell, K. *et al.* (1964a). A study of some cognitive and other disabilities in backward readers of average intelligence as assessed by a non-verbal test. *Brit. J. Educ. Psychol.*, 34, 58-64.

Lovell, K., *et al.* (1964b). A further study of some cognitive and other disabilities in backward readers of average non-verbal reasoning scores. *Brit. J. Educ. Psychol.*, 34, 275-279.

Lovell, K. (1966). The aetiology of reading failure. In *The First International Reading Symposium, Oxford*, 1964. London: Cassell, pp. 127-147.

McCarthy, J. J. and Kirk, S. A. (1961). *Illinois Test of Psycholinguistic Abilities.* Illinois: University of Illinois.

Monroe, M. (1932). *Children Who Cannot Read.* Chicago: University of Chicago Press.

Morris, J. (1966). *Standards and Progress in Reading.* Slough: National Foundation for Educational Research.

Myklebust, H. R. (1960). *The Psychology of Deafness.* New York: Greene and Stratton.

Prechtl, H. F. R., and Stemmer, C. J. (1962). The Choreiform Syndrome in children. *Develop. Med. and Child Neurol.*, 4, 119-127.

RUDNICK, M., STERRITT, G. M., and FLAX, M. (1967). Auditory and visual rhythm perception and reading ability. *Child Developm.*, 38, 581-588.

SEMMES, T., WEINSTEIN, S., GHENT, L., and TEUBER, H. L. (1955). Spatial orientation in man after cerebral lesion. *J. of Psychol.*, 39, 237-244.

SHAPIRO, M. B., BRIERLEY, J., SLATER, P., and BEECH, H. R. (1962). Experimental studies of perceptual anomaly VII. *J. Ment. sci.*, 108, 655-668.

STOTT, D. H., MAYES, F. A., and HEADRIDGE, S. E. (1966). *Test of Motor Impairment* Ontario : University of Guelph.

WEPMAN, J. M., and JONES, L. V. (1961). *The Language Modalities Test for Aphasia.* University of Chicago Educational Industries Service.

WERNER, E. E., SIMONIAN, K., and SMITH, R. S. (1967). Reading achievements, language functioning and perceptual motor developments of 10 and 11-year-olds. *Percept. and Motor Skills*, 25, 409-420.

(*Manuscript received 8th December*, 1967)

PSYCHOLINGUISTIC DEVELOPMENT AND THE CONTROL OF BEHAVIOUR

BY DENIS JOYNT
(The University of Papua and New Guinea)

AND BRIAN CAMBOURNE
(University College of Townsville)

SUMMARY. Luria maintains that one of the roles of speech in human development is to regulate behaviour, and that in the course of development this function passes through certain sequential stages that may be identified. The present investigation replicates Luria's study and is further concerned with the relationship between psycholinguistic abilities as measured by the Illinois Test of Psycholinguistic Abilities and the regulation of behaviour. The results show that Australian children pass through similar stages to Russian children in the development of the regulatory role of speech. In addition, these stages are more closely related to psycholinguistic development as measured by the ITPA, than to chronological age as Luria suggests.

I.—INTRODUCTION.

THERE is a large volume of research into various aspects of speech and language development. The work of Piaget (1955), Vygotsky (1962), and Klein (1966) shows that the majority of pre-school children talk to themselves, often in a manner that suggests that such speech has a definite orienting, organising and regulating function insofar as overt behaviour is concerned. The findings of the Kendlers (1962), Kuenne (1946) and Reese (1961, 1962) suggest that there is an important relationship between age, language development, and efficiency of verbally mediated problem solving behaviour. Pavlov's concept of the second signalling system has given impetus to research into semantic conditioning and concept formation. Results from such researches (Markosyan 1957, Luria and Vinagradova 1959, and others) show that when words are considered from the phonetic and semantic aspects, mature and undamaged nervous systems respond to the semantic aspects of words while immature or impaired nervous systems repond to the phonetic aspects.

In a series of investigations Luria (1957, 1959a, 1959b, 1960, 1961) has attempted to delineate clearly the developmental levels through which the child passes when mastering self-control. Based on his observations Luria proposes a sequence of four distinct stages in the development of the regulatory role of speech. He examines these stages under laboratory conditions. Working on the premise that the accomplishment of a simple action on verbal instruction is the core of verbally mediated behaviour, he used a relatively simple piece of equipment to analyse these stages of development. It consisted of a display panel for the presentation of stimuli, a rubber bulb that the subject squeezes to signify his responses, and a device to record both the stimuli presented and the changes in pressure on the bulb. The subject was expected to make a simple motor act (squeezing the bulb) in response to certain verbal instructions initially issued by the experimenter but later by the child himself. The difficulty of the tasks was altered either by varying the instructions or by making the stimuli more complex.

Luria's claims are consistent with the findings of American psychologists.

(*a*) He proposes a four-stage sequence of development.

Stage 1 ($1\frac{1}{2}/2$—$2\frac{1}{2}/3$ *years*) : The child's speech is insufficiently developed to serve as a regulator of his motor actions. He may be able to speak, pronounce words and understand their meanings but he is unable to use speech to direct his own behaviour. The speech of others, however, can direct and control his behaviour to some extent.

Stage 2 (3—4 *years*) : The role of external cybernetic signals is assumed by the child's own speech with the result that he can use his speech to control his behaviour to some extent. This is possible only when his self-instructions are phonetically and semantically congruent. Luria suggests that each phonetic element of a word serves to initiate motor movement in the young child. This impellent function may, or may not, coincide with the meaning of the word. For example, SQUEEZE has one impellent motor element and also means one motor movement. When, as in this case, the meaning of the word and its motor impellent function coincide, an instruction is considered phonetically and semantically congruent. On the other hand an instruction is considered phonetically and semantically incongruent when the impellent motor function(s) of the word or words conflict with meaning. TWICE, for example, in the experimental design *means* two squeezes but has one motor impulsive element only. Another example is DON'T SQUEEZE which has two motor impulsive elements but means inhibit motor movement. (See Appendix.)

Stage 3 (4—$5\frac{1}{2}$ *years*). The semantic (or meaningful) aspect of the child's speech begins to become dominant. He is able to respond to the meaningful content of his own spoken self-instructions, that is, even when the instructions are phonetically and semantically incongruent.

Stage 4 ($5\frac{1}{2}$ *years on*) : The child no longer audibly expresses self-instructions but silently produces behaviour which appears to be regulated by 'rules' formulated internally. This stage, Luria claims, is the ultimate stage of 'highest self-regulation' and, as such, is 'specifically human.'

(*b*) Progression from stage to stage results in increasing ability to regulate behaviour.

(*c*) The phonetic aspect of speech predominates in the early stages.

(*d*) The semantic aspect predominates in later stages.

(*e*) Progression from earlier stages to later depends on chronological development, that is, the improvement in language abilities which usually accompanies increase in chronological age.

Little or no work of a similar nature has been done with Australian children. Here, a major objective is to replicate Luria's studies using Australian subjects. The investigation is based on the hypotheses that Australian children pass through a series of stages in developing the regulatory role of speech similar to those described by Luria for Russian children, and that these stages are a function of the child's general psycholinguistic development, rather than chronological development as suggested by Luria.

II.—METHOD AND PROCEDURE.

Testing the hypotheses involves three steps :

(*a*) replication of the methods that Luria (1961) describes for investigating the sequential stages in the development of the regulatory role of speech ;

(*b*) comparison of the results of such a replication with the results that Luria obtained ; and

(*c*) comparison of the same results with the scores on the Illinois Test of Psycholinguistic Abilities.

Replication of Luria's Investigations.

Replicating Luria's methods presents a number of difficulties. To begin with no details concerning the presentation of stimuli are given. Although his apparatus is described at some length many important facts such as this are omitted from translations of his writings. Instead, narrative accounts of some of the subjects' behaviour is presented, illustrated with samples of raw data in the form of responses to stimuli presented by the experimenter. It is not known how Luria times the presentation of stimuli, or whether he uses a constant pattern of stimulus presentation in all experiments. In this investigation an average of 1·0 seconds for both intervals is used with a constant ratio of six positive to three negative signals. The order of presentation is based on the assumption that he would present his tasks in ascending order of difficulty.

Because of other omissions it has been necessary to make a number of assumptions about Luria's experimental design.

(*a*) The sample performed the tasks only once, so that there was no practice effect.

(*b*) The children were free from intellectual or sensory handicap and their language development appeared commensurate with their age.

(*c*) Each subject was treated identically. Rapport, incentives, verbal instructions, number of trials, physical arrangements, and so on, were similar for each child.

Three other points should be noted in connection with the experimental design. In the first place, the child's whispered commands to himself appear to be less effective than louder commands. Although Luria did not say whether this held for all ages his theory seems to predict that the effect should vary with the subject's age. Secondly, it appears that the child's speech sometimes precedes and sometimes follows the motor behaviour it is said to be directing. Luria (1960b) merely noted the existence of this pheomenon. Thirdly, he is vague about the time lapse between attainment of stages three and four. On some occasions (1961, p. 59) he seems to imply that Stages 3 and 4 may be attained *simultaneously*, the attainment of one presupposing the other. However on other occasions (1961, p. 61) he states clearly that Stage 4 *follows* Stage 3 and that there is a period during which the child, though capable of responding to the semantic aspects of words, still cannot direct his behaviour silently. Such apparent contradictions are serious obstacles to a thorough understanding of the developmental sequence Luria proposes.

Comparison with Luria's Results.

Comparison of the results of this investigation with Luria's also presents a number of difficulties. One is the selection of a sample of comparable Australian children. Luria's sample is described as ' normal ' but no criteria of normality are offered. Here, children in whom intellectual or sensory handicap is obvious, are excluded. However, it is impossible to be certain that subjects with minimal handicaps or some degree of language deficiency were also excluded.

Performance on Luria's Tasks and Psycholinguistic Development.

The Australian subjects' language development is expressed in terms of a standardised language age derived from performance on the Illinois Test of Psycholinguistic Abilities (ITPA). Although probably culturally biassed to some extent, the test does provide the best available means of grading the sample in terms of language development. In addition, evaluation of subtest scores may help isolate and measure some of the factors involved in the verbal mediation of the behaviour necessary to perform Luria's sensorimotor tasks. (Analysis of these data will be published elsewhere). In this study a comparison is made between the score on Luria's tasks and the stage of language development as measured by the ITPA.

Presentation of Results.

Before any comparisons may be made, an objective method of comparing the two sets of results must be devised. Therefore a criterion of success or failure on the sensorimotor tasks is essential. Luria (1961) describes his criterion of success as a correspondence between verbal instruction and motor response, indicated in all situations by a disappearance of intersignal pressures of the bulb and the appearance of clear cut reactions co-ordinated with the conditioned signal. It is not stated explicitly how clear cut these reactions are to be or what degree of co-ordination is meant.

A pilot study revealed that success or failure on any of the sensorimotor tasks is not an all-or-nothing matter, but rather than there are many degrees of success or failure. Luria expressed success in terms of a diminishing number of errors. Here an Error Index and a category of Disconnected Behaviour are used to refine the process of measurement.

In order to allow comparison to be made between responses to the different numbers of stimuli that were presented on a few occasions, an Error Index is used in place of raw scores. The Error Index is the ratio of total errors to total possible correct responses on any single task. Increase in size of the index represents greater failure.

Use of the Error Index as a basis for scoring masks differences between types of error. To compensate for this use is made of the fact that it is possible to recognise from the child's behaviour when an error occurs because a particular sensorimotor task is beyond his capabilities. Thus, when the subject is unable to connect the stimulus presented with either the verbal instruction or the required action, this is termed Disconnected Behaviour.

There are two criteria of Disconnected Behaviour: firstly, when subjects are engaged on a task under conditions where self-administered instructions are given and they squeeze indiscriminately and impulsively without using audible speech as directed, or fail to squeeze but use audible speech as directed; and secondly, when engaged on a task under the silent condition, the subject squeezes impulsively in an indiscriminate manner, regardless of the appearance or disappearance of stimulus signals. In this case the subject usually gazes around the room or at the bulb, apparently unable to co-ordinate his behaviour with the demands of the task. The category of Disconnected Behaviour allows more accurate assignment of subjects to the appropriate level on Luria's developmental tasks.

Subjects.

Thirty-two subjects, ranging in age from 1 year 6 months to 7 years 7 months, were chosen on the basis of availability and apparent freedom from sensory, speech, or intellectual defect.

Apparatus.

The apparatus is designed to present visual stimuli manually from a point behind the subject in such a way as to preclude anticipation of the impending stimulus, and to record presentations of stimuli and response on a moving tape.

Experimental Conditions.

Reports of Luria's work show that he conducted his investigations using variations of certain sensorimotor tasks to ascertain the subject's stage of development. Some variations he describes as 'more complex' than others (1961, p. 72). Here four comparable experimental conditions are adopted.

(*a*) *Experimenter's Speech Only Condition* (*ESO*).

In the ESO condition verbal instructions to regulate bulb squeezing behaviour are issued only by the experimenter. The child is instructed to start squeezing the bulb and, once started, is told to stop.

(*b*) *Subject's Audible Speech Condition* (*SAS*).

In the SAS condition the child audibly regulates his own bulb-squeezing behaviour in two different ways.

(i) SAS—A : Here the words used in the self-administered instructions are phonetically and semantically congruent.

(ii) SAS—B : Here the words used are *not* phonetically and semantically congruent.

(*c*) *Silent Condition* (*X*).

In the silent X condition no audible speech is used by children when attempting the tasks.

In each of the SAS conditions and in the X condition the following three tasks are used :

Number 1—The red light is flashed on and off regularly. The child is directed to squeeze the bulb every time the red light appears.

Number 2—The child is shown a pattern of red and green lights. He is directed to squeeze for the red but not for the green. The pattern shown is three red, one green, two red, one green, one red, one green.

Number 3—The red light is flashed on and off regularly. The child is directed to make two responses to each signal.

Implicit in translations of Luria's work are the following assumptions concerning these sensorimotor tasks.

(*a*) The most complex of the four conditions used is the X condition, because a child who successfully carries out the tasks silently has reached the ultimate stage of the proposed sequence.

(*b*) Similarly, the SAS—B condition, which utilises verbalisations which are not phonetically and semantically congruent to regulate motor behaviour, shows that Stage 3 has been reached.

(*c*) Success in the SAS—A condition, which uses verbalisations that are phonetically and semantically congruent, shows that Stage 2 has been reached.

(*d*) Success in the ESO condition means that the child's speech is at Stage 1.

Although ability to complete successfully the sensorimotor tasks in the silent condition presupposes sufficient ability to succeed at lower tasks, all subjects were tested in all conditions. Each commenced with the ESO condition, followed by SAS—A, SAS—B, and X. Within three days of this procedure being completed the ITPA was administered.

Procedure.

To compare Luria's (1961) results with the findings of this study, the following procedure was adopted.

(*a*) Subjects were arbitrarily divided into four equal groups on the basis of chronological age. Thus the youngest became Group 1, the next youngest Group 2, and so on for Groups 3 and 4, (*b*) The Error Index for each subject in each group was calculated for all sensorimotor tasks.

(*c*) The Mean Error Index for each group was calculated for all sensorimotor tasks. These means are compared.

(*d*) Subjects were similarly divided on the basis of language age derived from the ITPA.

(*e*) The relationship between chronological age and sensorimotor performance was calculated.

(*f*) The relationship between language age and sensorimotor performance was calculated similarly.

(*g*) The significance of the difference between the coefficients for both methods of grouping was calculated.

For a detailed description of instructions given to subjects, see Appendix.

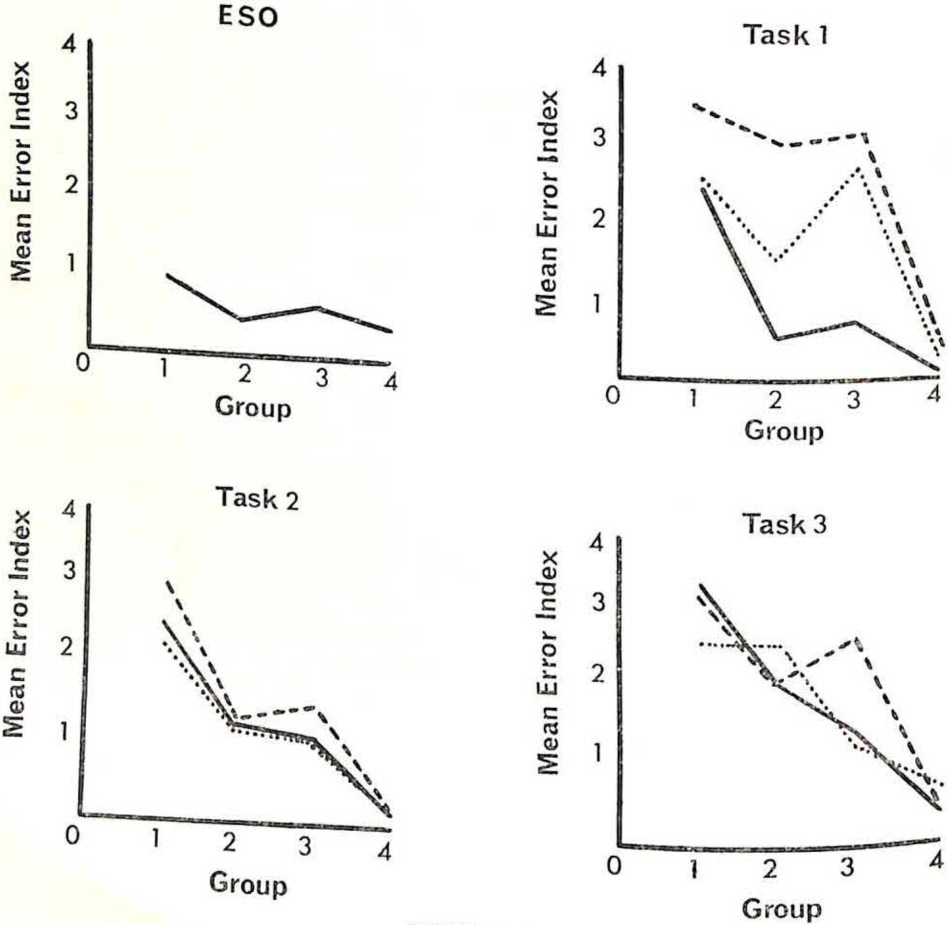


FIGURE 1

Comparison Mean Error Indices for Chronological Age Groups.

——— A Condition : Speeech-Motor Activity Congruent.

- - - - - - B Condition : Speech-Motor Activity Incongruent.

··········· X Condition : No Audible Speech.

III.—RESULTS.

When subjects are grouped by chronological age and the results are compared (see Figure 1) there is a tendency for mean errors to decrease as the age of the groups increases. However this is a general tendency only as on six occasions Group 2, which has a mean chronological age of some ten months less than Group 3, has a lower incidence of errors.

When subjects are grouped according to language age and their results compared (see Figure 2) a definite pattern emerges. In every instance there is a progressive decline in mean errors as the mean language age of the group increases.

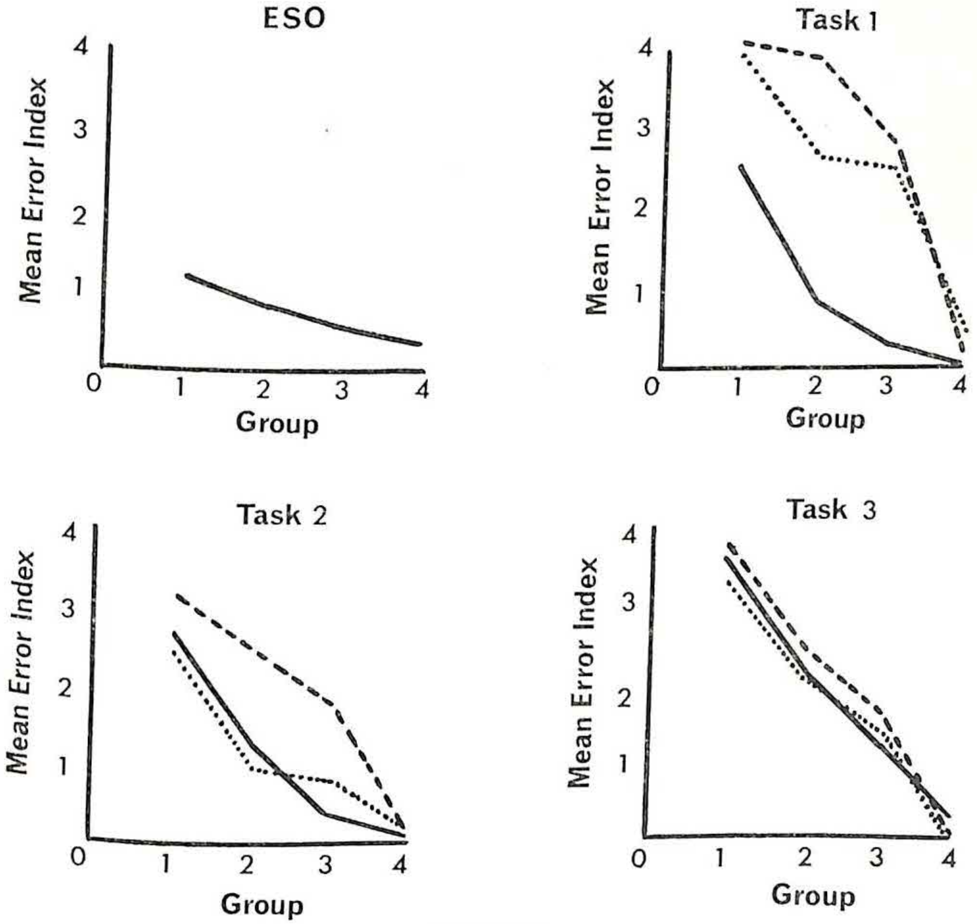


FIGURE 2

COMPARISON MEAN ERROR INDICES FOR LANGUAGE AGE GROUPS.

—— A Condition : Speech-Motor Activity Congruent.

- - - - B Condition : Speech-Motor Activity Incongruent.

·········· X Condition : No Audible Speech.

Apart from extreme cases it is difficult to determine specifically to which of Luria's four stages of development an individual child properly belongs. All children who attain a language age of 88 months or more show very clearly that they have reached Stage 4. Similarly subjects whose language age is 38 months or less show just as clearly that their development corresponds with Stage 1. Although the latter's error indices under Stage 2, 3 and 4 conditions do not appear numerically much greater than some psycholinguistically older children they represent a grossly different type of behaviour. The errors consist predominantly of impulsive pressures of the bulb, totally unrelated to the stimuli presented or the speech used.

Other subjects are more diffcult. The majority make some *apparently* correct responses under the four testing conditions. The responses are described as apparently correct because it is virtually impossible to decide objectively whether a success was by chance or the result of some form of mediated behaviour. In such circumstances no final decisions may be made about the stage of development a child's speech has reached in its role as a regulator of behaviour.

Because of such problems it is difficult to analyse with any degree of precision the relationship between Luria's proposed stages of development and language age. However, if the assumption underlying the categorisation of disconnected behaviour is accepted, then the following picture appears.

A language age of 29 months shows that language as a regulator of behaviour is in Stage 1 of Luria's proposed sequence of development. In this study this basal age is slightly inflated as the youngest children in the sample were not able to cope with the lowest level of the ITPA (30 months) and were arbitrarily assigned to language ages of 29 months. Stage 1 is characterised by language ages up to 43 months, when transition to the second stage appears to be taking place.

In the present sample one subject with a language age of 43 months is judged to be displaying disconnected behaviour at Stage 2 of Luria's sequence. Another, of identical language age, does not do so and shows the ability to make some correct responses at the Stage 2 level of performance. Such results suggest that an Australian child with a language age of 43 months has, or is about to, progress to Stage 2. This stage persists, with gradual improvement, until a language age of about 51 months is reached.

Stage 3 is apparently attained by children ranging in language age from 54 to 64 months although overlapping with the stages above and below is noticeable. Nevertheless overlap of more than three months is rare. Children who have attained a language age of 61 months or more are characteristic of Stage 4. However, no definite statement can be made about progression from Stage 3 to Stage 4. As has been pointed out Luria is also vague about this. Our results show that some Australian children of language age 54–64 months are capable of regulating their behaviour in response to the semantic aspects of speech used externally, while incapable of regulating the same behaviour in response to silent internal speech. Others can successfully regulate their behaviour using both external and silent speech at about 61 months of age. Apart from longitudinal studies there does not appear any way of discovering whether both skills are sometimes attained simultaneously. We believe that some Australian children do attain both almost together, while others remain in the third stage for some three or four months. The relationship between language age and stage of development is shown in Table 1.

TABLE 1

SUBJECTS JUDGED TO MANIFEST DISCONNECTED BEHAVIOUR UNDER VARIOUS TESTING CONDITIONS.

Condition	Name	L.A. (in mths.)	Stage of Luria's Proposed Sequence
ESO All Subjects were capable of performing this task in a co-ordinated, connected manner.			
SAS–A	Cathy K.	29	Stage 1 Language sufficiently developed to respond to both initiating and inhibiting verbal instructions from another person. Insufficiently developed to direct own behaviour using own speech.
	Brooke S.	29	
	Margo M.	29	
	Brock C.	29	
SAS–B	Peter T.	29	
	Wendy N.	39	
	Tim B.	40	
X	Derry N.	38	
	Pedro M.	41	
	Cathy M.	43	
SAS–B	Grahame C.	46	Stage 2 Language sufficiently developed to enable subject to regulate behaviour using audible self instructions providing that the semantic and phonetic aspects of the verbalisations are congruent with the motor actions. Insufficiently developed to use instructions comprising verbalisations which are semantically and phonetically incongruent.
	Elizabeth G.	49	
	Susan P.	57	
	Megan W.	51	
	Geoffrey R.	43	
X	Fay F.	46	
X	Marylin L.	54	Stage 3 Language sufficiently developed to regulate behaviour using verbalisations not semantically and phonetically congruent. Insufficiently developed for silent self administered directions.
	Kate M.	53	
	Lyn S.	64	
	Mathew T.	64	
	Chris. M.	63	
	Stephen R.	55	
No disconnected behaviour on any of testing conditions.	Stephen D.	61	Stage 4 Language sufficiently developed to regulate behaviour using internal self administered directions.
	Simon V.	65	
	Noel H.	50	
	Scott D.	67	
	Mark G.	73	
	Michael B.	70	
	John E.	69	
	Judy B.	88	
	Kay A.	90	
	Brian N.	99	

In one exceptional instance a subject with a language age of 50 months performed the tasks with some degree of success at all levels. This subject constantly gave himself instructions during the ITPA testing, even when told not to say anything.

Table 2 gives the correlation coefficients showing the degree of relationship between chronological age and sensorimotor performance and between language age and sensorimotor performance. (The correlation between chronological age and language age for this sample is 0·90).

TABLE 2

RANK CORRELATIONS OF ERROR INDICES WITH CHRONOLOGICAL AGE AND LANGUAGE AGE.

Condition/Task		Correlation of Error Indices with		Significance of Difference
		Chronological Age	Language Age	
ESO		·45	·58	·05
SAS–A	Task 1	·80	·87	N.S.
	2	·80	·90	·01
	3	·85	·88	N.S.
SAS–B	Task 1	·68	·85	·01
	2	·81	·94	·01
	3	·80	·90	·01
X	Task 1	·71	·76	N.S.
	2	·83	·88	N.S.
	3	·79	·87	·01

In the first case the correlations are greater than 0·70 in all but two instances, suggesting a general tendency for increase in chronological age to be accompanied by an increase in the ability to regulate verbally some forms of sensorimotor behaviour. In the second case, the correlation coefficients are greater than 0·85 in all but two instances. This indicates a much closer relaship, six of the ten pairs of correlations having a significant difference at the 0·01 level. This suggests that Australian children's ability to verbally regulate behaviour on the tasks devised by Luria is more a function of psycholinguistic development as measured by the ITPA than of the general development that accompanies increased chronological age. It seems that the ITPA battery of tests and Luria's sensorimotor tasks sample very similar abilities.

IV.—DISCUSSION.

When Australian and Russian samples are compared they show similar characteristics in the relationship between chronological development and gradual increase in the power to regulate behaviour. More important, perhaps, is that an even more significant relationship is shown in the Australian sample, between children's psycholinguistic development and certain aspects of self-control. When correlated with language age these results indicate four stages of development. In Stage 1, ranging in language age up to about 43 months, speech becomes sufficiently developed to respond to both initiating and inhibiting verbal instructions from another person, but is insufficiently developed for the child to use his own speech to direct his behaviour. Stage 2, which covers a language age from approximately 43 to 51 months, is characterised by speech sufficiently developed to enable the child to regulate his behaviour using audible self-instructions, provided that both the semantic and phonetic aspects of the verbalisations are congruent with the motor actions. However, speech is insufficiently developed for instructions that are semantically and phonetically incongruent to succeed. The third stage, in which the language age ranges from 54 to 64 months approximately, is marked by speech that is developed sufficiently to regulate behaviour using verbalisations that are not phonetically and semantically congruent, but insufficiently developed for silent self-administered directions to be successful. In the final stage, from about 65 months upwards, speech is sufficiently developed to regulate behaviour using silent self-administered instructions.

The relationship between psycholinguistic development and the ability to regulate behaviour through speech throws a new light on the importance of encouraging early speech in children. Promotion of such language development may lead to the child attaining more rapid control over some aspects at least of his own behaviour, and hence to more efficient control of his environment. However, more refined analysis of psycholinguistic processes and development is necessary before specific recommendations may be made. In addition, measurement of the child's psycholinguistic development gives an indication of his functioning ability to regulate his own behaviour. Thus the demonstrated relationship suggests the opportunity to facilitate such control and offers ways of measuring its growth.

In conclusion, it seems that Luria's investigations, limited though they are in many ways, open a new area of enquiry which had great relevance for researchers in the general field of cognitive development but, more particularly, shows promise for those concerned with the training and education of children.

V.—REFERENCES.

KENDLER, T. S., and KENDLER, H. H. (1962). Mediated response to size and brightness as a function of age. *Amer. J. Psychol.*, 75, 571-586.

KLEIN, W. K. (1966). An investigation of the spontaneous speech of children during problem solving. Cited by S. E. Ervin-Tripp and D. I. Slobin in Psycholinguistics, *Annual Review of Psychology*.

KUENNE, M. R. (1946). Experimental investigation of the relation of language to transposition behaviour in young children. *J. Exp. Psychol.*, 36, 471-490.

LURIA, A. R. (1957). The role of language in the formation of temporary connections. In B. Simon (ed.) *Psychology in the Soviet Union*. Stanford : Stanford University Press.

LURIA, A. R. (1959a). The directive function of speech in development and dissolution, 1—Development of directive function of speech in early childhood. *Word*, 15, 341-342.

LURIA, A. R. (1959b). Experimental analysis of the development of voluntary action in children. In *The Central Nervous System and Behaviour, Selected Translations from the Russian Medical Literature*. U.S. Department of Health, Education and Welfare.

LURIA, A. R. (1960). *The Nature of Human Conflict*. New York : Grove Press.

LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour*. London : Pergamon.

LURIA, A. R., and VINAGRADOVA, O. S. (1959). An objective investigation of the dynamics of semantic systems. *Brit. J. Psychol.*, 50, 89-95.

MARKOSYAN, A. A. (1957). On the interaction of two signal systems in involuntary processes. In *Proceedings of Conference on Psychology*. Moscow : Pedagogical Academy of Sciences.

PIAGET, J. (1955). *The Language and Thought of the Child*. New York : Meridian.

REESE, H. W. (1961). Transposition in the intermediate size problem by pre-school children. *Child Develop.*, 32, 311-314.

REESE, H. W. (1962). Verbal mediation as a function of age level. *Psychol. Bull.*, 59, 502-509.

VYGOTSKY, L. S. (1962). *Thought and Language*. Cambridge : Massachusetts Institute of Technology Press.

(*Manuscript received 30th January*, 1968)

VI.—APPENDIX.

PROCEDURE.

The instructions issued for each of the testing situations were :

(*a*) *ESO Condition*.

See the rubber ball ? When I tell you I want you to pick it up in your hand and start squeezing it, like this. (*E* demonstrates.) Keep squeezing it without stopping. (*S* is allowed to squeeze for 5 to 10 seconds.) Stop, don't squeeze any more. I want you to stop now.

These instructions were given several times to each *S*.

(*b*) *SAS Condition—Part A* (*SAS—A*).

Task Number 1: Everytime the red light comes on, tell yourself SQUEEZE (or NOW) so that I can hear you, and then squeeze the ball—like this. (*E* demonstrates.)*

Task Number 2: Everytime the red light comes on, tell yourself SQUEEZE (or NOW) so that I can hear you, and squeeze the ball. Everytime the green light comes on just watch it. Don't say anything, don't squeeze the ball, but watch the light like this. (*E* demonstrates.)

Task Number 3: Do you know how many two is? Clap your hands two times. Stamp your feet two times. Show me two fingers. Squeeze the ball two times. (*E* observed and assisted *S*s until each indicated that a concept of 'two times' was understood.) Everytime the red light comes on tell yourself SQUEEZE—SQUEEZE (or NOW—NOW) so that I can hear you, and squeeze the ball two times, like this. (*E* demonstrates.)

(*c*) *SAS Sondition—Part B* (*SAS—B*).

Task Number 1: Everytime the red light comes on tell yourself I MUST SQUEEZE NOW so that I can hear you, and then squeeze the ball—like this. (*E* demonstrates.)

Task Number 2: Everytime the red light comes on tell yourself SQUEEZE so that I can hear you, and squeeze the ball. Everytime the green light comes on tell yourself DON'T SQUEEZE so that I can hear you and don't squeeze the ball, just look at the light. (*E* demonstrates.)

Task Number 3: Do you know the grown-up word for two times? It is twice. What is it? Clap your hands twice. Stamp your foot twice. Squeeze the ball twice. (*E* observes and assists where necessary.) Everytime the red light comes on tell yourself TWICE so that I can hear you, and squeeze the ball twice—like this. (*E* demonstrates.)

(*d*) *X Condition*.

Task Number 1: Everytime the red light comes on squeeze the ball. Don't say anything, just squeeze the ball. (*E* demonstrates.)

Task Number 2: Everytime the red light comes on squeeze the ball. Don't say anything, just squeeze the ball. Everytime the green light comes on just watch it. Don't say anything and don't squeeze the ball, just watch the green light. (*E* demonstrates.)

Task Number 3: Everytime the red light comes on, squeeze the ball two times. Don't say anything just squeeze the ball two-times. (*E* demonstrates.) It should be noted that *E* demonstrates by performing the task himself. He also repeats the instructions several times and, if *S* is mature enough, asks him to repeat them.

---

* If some of the very immature children could not pronounce SQUEEZE, a simpler word with similar phonetic-semantic characteristics (e.g., NOW) was substituted.

THE FORMATION AND EXTENSION OF CLASS CONCEPTS IN PRIMARY SCHOOL CHILDREN

By E. STONES and J. R. HESLOP
(*School of Education, University of Birmingham*)

Summary. The aims of the investigation were (1) to establish the extent to which Vigotsky's model of concept formation applied to English primary school children; (2) to examine the relationships between age, verbal ability, non-verbal ability and levels of thinking; and (3) to examine the relationship between level of thinking and performance in tests of transfer. 60 primary school children aged 6 years to 11 years were given the Vigotsky sorting task and specially constructed tests of extension (or transfer). It was found that the Vigotsky model was appropriate although the subjects showed little evidence of thinking at the lowest level. Verbal ability was associated more closely with conceptual thinking than the other measures. Performance on tests of transfer was highly correlated with level of thinking in the concept forming situation.

I.—Introduction.

The appearance of the translation of Vigotsky's book on language and thinking (1962, originally published in Russian in 1934) has stimulated a new interest in the inter-relationships of language and concept formation. In his book Vigotsky reports work which investigated these inter-relationships. On the basis of this and other work he concluded that language has a profound influence on the formation of concepts. He considered that concept formation cannot be reduced to " . . . association, attention, imagery, inference, or determining tendencies. They are all indispensable but they are insufficient without the use of the sign, or word as the means by which we direct our mental operations, control their course, and channel them towards the solution of the problem confronting us."

The difficulty in investigating the influence of language on concept formation is that words normally symbolise concepts. Thus any words used in an experiment would already have conceptual content and would therefore introduce error. Vigotsky overcame this difficulty by using an adaptation of the apparatus devised by Ach and described by Hanfmann and Kasanin (1937).

The apparatus which Vigotsky used comprises twenty-two small wooden blocks of five different colours, six different shapes of cross section, two different heights or thicknesses, and two different cross sectional areas. Under each block is written one of four nonsense words: LAG, BIK, MUR and CEV. The problem is to classify the blocks in four groups so that all the blocks in any one group have common properties which unequivocally mark them off as members of that group and as non-members of any other group. The criterial attributes for 'correct' grouping are cross-sectional area and height. The nonsense syllables relate to the criterial attributes and may, in fact, be considered as 'names' of the concepts embodied in the blocks. Thus, the LAG blocks are tall and fat, the BIK blocks are small and fat, the CEV blocks are small and thin, and the MUR blocks are tall and thin.

This apparatus presents the subject with two sets of stimuli; one set, the blocks themselves, and the other, the nonsense words. The blocks provide the object of the subject's activity and the nonsense words can serve to organise his activity. Thus a subject is faced with a task which will be solved when he has acquired new concepts for which the language has no names. The nonsense words are no help in the beginning since they have no conceptual content, but if the subject uses the words to organise and direct his activity they will be increasingly useful as the experiment proceeds and will finally come to symbolise the newly acquired concepts. They will no longer be nonsense words to the subject but will, in fact, be the 'names' of the concepts. One other feature of the apparatus is that it enables the experimenter to observe the steps taken in the formation of new concepts since every move involves the manipulation of the apparatus and is thus externalised and apparent.

The Vigotsky blocks have been employed by psychologists for many years but not always in the way Vigotsky used them. Semeonoff used them as part of a testing programme for 'special service' work during war time. The apparatus was later developed and standardised by Semeonoff and Laird as a test of intelligence (Semeonoff and Laird 1952). Semeonoff and Trist (1958) used the apparatus as a test of performance. Hanfmann and Kasanin (1937) used the test in the study of concept formation and as an instrument in the study of schizophrenia. The history of the development of the apparatus and a detailed account of its rationale are given by Hanfmann and Kasanin.

Vigotsky used the blocks with normal and with schizophrenic subjects. His book refers particularly to his work on concept formation with normal subjects including children and this is where our interest lay. In his experiments Vigotsky identified three main phases in the formation of concepts. The first is syncretic thinking where the blocks are grouped almost fortuitously, either randomly or according to criteria such as accidental contiguity. The second phase is thinking in complexes in which the subject makes use of bonds actually existing between the experimental objects. These bonds are concrete and factual rather than abstract and logical. Vigotsky found several types of complexes which he discusses in detail in his book. The third phase begins when the subject begins to abstract the qualities of the experimental material. In this phase he is able to form *potential concepts*. Potential concepts abstract only a single attribute of the experimental objects. It is often difficult to differentiate clearly between this mode of grouping and pseudo-concepts which are a form of complexive grouping. The main difference between pseudo-concepts and potential concepts is seen when the subject is shown that his grouping is incorrect. The subject who is thinking conceptually will have formed a hypothesis and will abandon it when proved wrong by being given a wrongly chosen object, and will try another method of grouping. The subject who is thinking in pseudo-concepts will maintain his original grouping and only remove the example which the experimenter shows to be wrong even though several other objects may also be wrong. Potential concepts are eventually replaced by true conceptual grouping when the subject classifies consistently according to the two criterial attributes.

Vigotsky stressed the importance of language in concept formation. Inhelder and Piaget (1964) considered that whilst classification is possible without language, the extensive aspect of classification is probably dependent upon fairly sophisticated use of language. Bruner and his associates (Bruner *et al.* 1966) in a series of experiments which owe much to Vigotsky's earlier work, investigated the inter-relationships between language and conceptual grouping. Using everyday articles, pictures and words in studies of classificatory ability, they found groupings very similar to those observed by Vigotsky. Their investigations did not study the formation of concepts but the type of conceptual groupings used by subjects of different levels of linguistic sophistication. Their findings indicate strongly that level of sophistication in classifying is closely related to level of sophistication in the use of language.

Annett (1959) in a sorting test which made use of material familiar to the children (pictures of animals, plants, vehicles and furniture) found that the sorting strategies of children went through stages which are not apparent on consideration of their final form in adults. She also found groupings which resembled those found by Vigotsky and Bruner. She did not specifically investigate the role of language.

Vigotsky's main interest was to investigate the process of concept formation as evidenced by the moves made in the sorting of the blocks. In particular, he was interested in the interaction of the verbal label and the manipulation of its concrete referents in the process of concept formation. This was also one of our key interests. We wished to investigate the applicability of Vigotsky's model to concept learning in English children of primary school age. We were also interested to investigate the relationship between modes of thinking and measures of non-verbal and verbal ability. We hypothesised, on Vigotsky's model, that there would be a closer relationship between verbal ability and mode of thinking than between non-verbal ability or age and mode of thinking. In one particular our interests went beyond the experiments reported by Vigotsky. Vigotsky gives little information about tests of transfer of the concepts once acquired, although he does refer to such tests. We considered this to be important and decided to investigate the extensive aspect of the concept formation task as well as the intensive aspect (Piaget and Inhelder 1957).

II.—PROCEDURE.

Apparatus.

The apparatus consisted of the standard Vigotsky blocks, six tests of 'extension' and tests of verbal and non-verbal ability.

The extension test apparatus consisted of materials which differed from the blocks in all particulars except the two criterial attributes of height and cross-sectional area. The first extension test made use of four plasticene men modelled to embody the qualities symbolised by the nonsense words on the blocks: i.e., there was a short fat man, a short thin man, a tall thin man and a tall fat man. The second test used four pictures of buildings drawn in perspective, according to the same principles as the material in test one. Test three used four pictures of trees drawn without perspective,. The material for test four was a piece of plasticene from which the children had to model exemplars of the four concepts. Test five, which needed no apparatus, is described below.

The Simplex Group Intelligence Scale and the Raven's Progressive Matrices were administered. Children in the 6–7-year group were unable to do the Simplex test and thus no results on this test were available for them.

Subjects.

The sample comprised 60 children randomly selected from a primary school in a high socio-economic urban community. Ten children were taken from second-year infants, third-year infants and from each of the first, second, third and fourth years of the junior school.

The classification task.

At this stage of the investigation there was no intention to use the apparatus as a test. Our main concern was, as with Vigotsky, to examine the kinds of classificatory moves made by the children. We were thus investigating the ways in which children learned new concepts and not testing their grasp of previously acquired concepts. The procedure followed was based on the method adopted by Hanfmann and Kasanin. The blocks are spread randomly and evenly over a board 25-in. × 20-in. with the nonsense words turned down. The experimenter explains that the blocks can be arranged in four groups and that each group has a name. The task is to arrange the blocks in their correct groups. The subject is shown an example by the experimenter's turning over one of the blocks and showing the nonsense word. The experimenter asks the subject to arrange the blocks in their correct groups and emphasises that he is free to move blocks from one group to another or back to the pool of unsorted blocks whenever he wishes. The subject is also instructed not to upturn any blocks and is encouraged to verbalise as he is grappling with the problem. This is intended to give added data on the method of attack on the problem. Whenever a subject completes a sorting incorrectly, the experimenter turns over one incorrect block, shows that it is different from the specimen and encourages the subject to try again. With each wrong sorting the number of upturned blocks increases and provides additional clues to the subject.

Vigotsky found that as the number of upturned blocks increases the subject obtains by degrees a basis for discovering to which characteristics of the blocks the nonsense words refer. When the subject makes this discovery the words come to stand for definite kinds of objects (e.g., LAG for large tall blocks, BIK for large flat ones), and " . . . new concepts for which the language provides no names are thus built up." When this stage has been reached the subject is able to complete the task with facility. This type of solution is achieved only by subjects who, in the sorting, eventually abstract the critical attributes of the concepts. It is, however, possible to complete the grouping without abstraction. The extreme case of this type of grouping would be when so many wrong groupings are made that almost all the blocks are turned up by the experimenter. The subject is then able to use the nonsense words as marks of identification and has merely to put the blocks with the appropriate words together. The nonsense words would in such a case have no symbolic qualities at all. As a first test to check on the degree of abstraction achieved by each child in the sorting, the blocks were brushed up after the first sorting was completed, the nonsense words turned down again, and the subject asked to repeat the sorting. It was considered that subjects who had acquired the new concepts and had abstracted their criterial attributes from the apparatus would be able to re-group the blocks correctly at once, whereas other subjects would have difficulty.

The extension tests.

The extension tests were administered after the grouping task had been completed. No blocks were visible at this stage of the experiment. In test one the children were shown the four plasticene men and asked to classify them according to the names used in the sorting phase. Tests two and three similarly demanded that the children classify the pictures as LAG, BIK, MUR, or CEV. In test four the children were each given a lump of plasticene and asked to make four models of Christmas parcels having the qualities of the four concepts. To score correctly the models had to show clearly the two distinctive attributes of the concepts.

In test five the subjects were asked the meaning of LAG, BIK, MUR and CEV. To score on this test the subjects had to refer to the two criterial attributes of the concepts.

Scoring.

In order to obtain an accurate record of the interviews, a method of coding children's responses and classifying moves was devised and a standard protocol was completed for every subject. In addition, a tape recording was made of every interview. Each experiment was conducted in a quiet room in the child's own school. Pilot studies were conducted to develop and refine the experimental procedure. There was no time limit and every effort was made to establish good rapport. A very detailed account of each interview was thus obtained and was later used to analyse the type of grouping the child was using. During the classificatory phase of the investigation no attempt was made to use the methods of scoring employed when the apparatus is used in testing. Our main concern was to discover the modes of grouping adopted by the children rather than to consider the task as a performance test. However, as a result of our analysis we were able to quantify the types of grouping used.

The extension tests were a different proposition. At this stage we were interested to see to what extent the new concepts had been learned. We therefore awarded one point for each correct classification in each of the first three tests. In test four we awarded one point for each model which clearly exemplified the two criterial attributes. Test five was scored by giving one point for each definition which clearly stated the two attributes of the concepts. Thus every test had a maximum score of four marks.

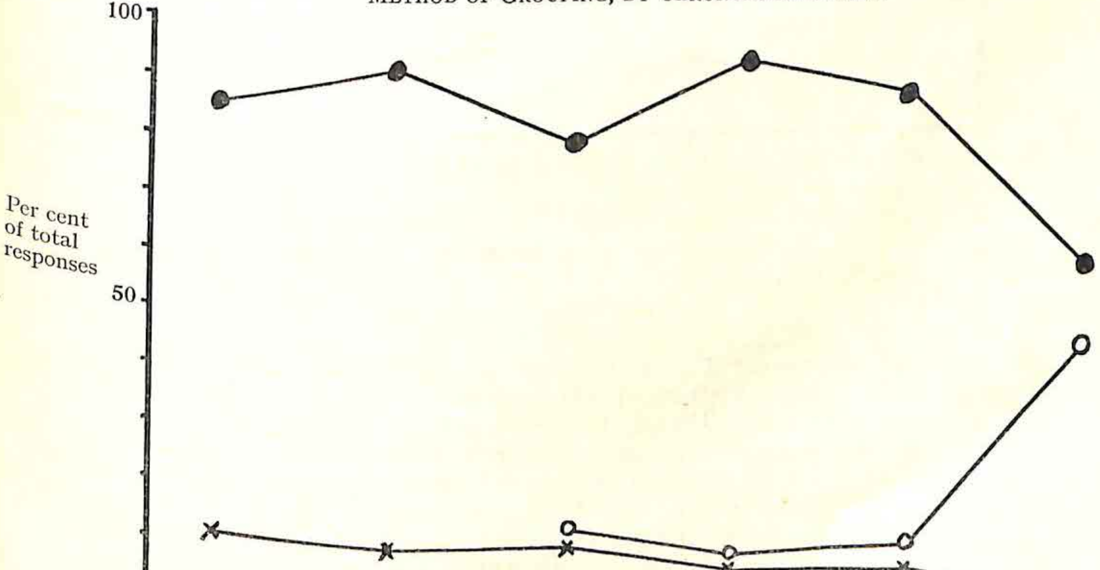
III.—RESULTS.

Types of grouping made.

The types of grouping found by Vigotsky were observed in the experiment. Within the three main phases of grouping (i.e., phase one, pre-complexive; phase two, complexive; and phase three, conceptual) we found examples of all the sub-phases described by previous investigators. We do not report these groupings in detail since the experiment is principally concerned with the three major phases of grouping activity. The responses of the children were classified according to the categories given and the proportion of each child's responses in the given phases was computed. The mean proportions of the responses in the three phases made by the ten children from each of the six (school) classes were then calculated to give the picture presented in figure one.

FIGURE 1

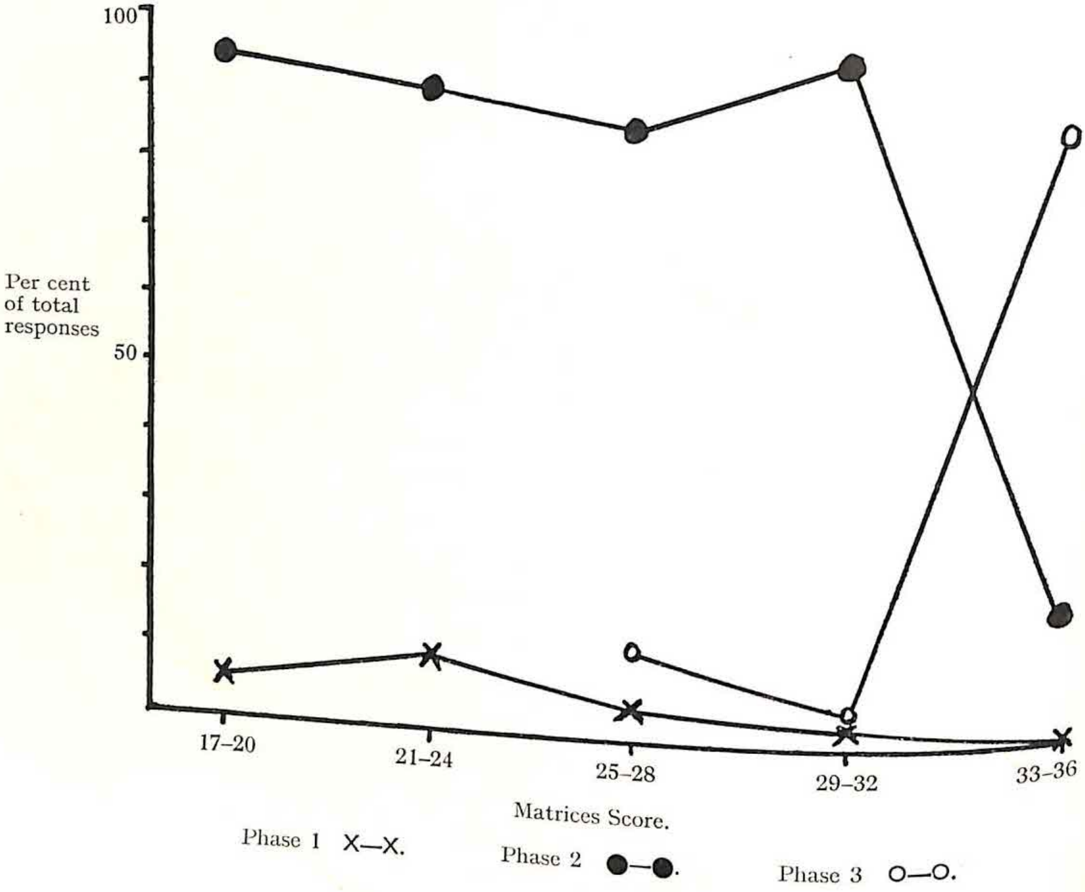
METHOD OF GROUPING, BY CHRONOLOGICAL AGE.



Rearranging the data to plot mean proportion of responses in the different phases against verbal and non-verbal ability we have the picture presented in Figures 2 and 3.

FIGURE 2

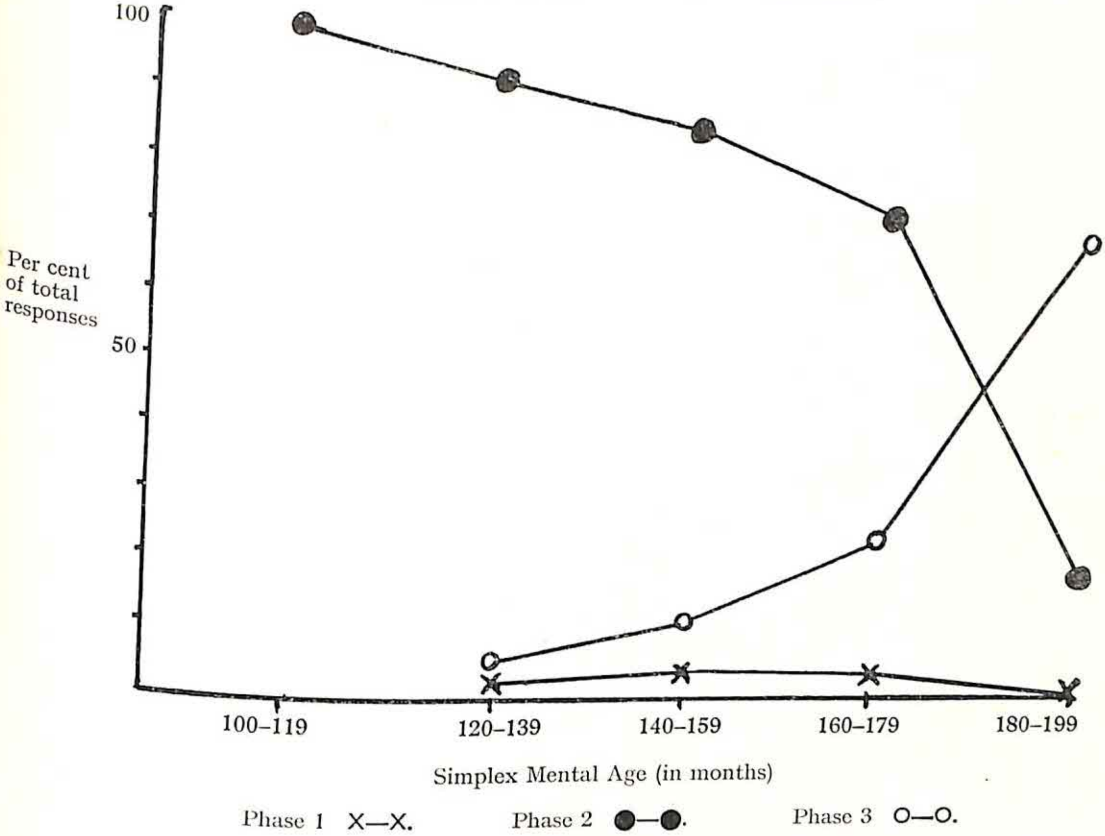
METHOD OF GROUPING, BY NON-VERBAL TEST SCORE.



Most subjects operated in more than one mode, grouping the blocks in different modes and combining and switching from one mode to another in the course of the experiment. Figures 1, 2 and 3 indicate the mean proportion of responses in each phase for each of the alternative methods of grouping. It is apparent from these figures that complexive grouping predominates in the sorting activity of this group of children. Random sorting, what Vigotsky called "vague syncretic conglomerations," is little in evidence with children in this sample, accounting for about 10 per cent of total responses at about age 6·6 and gradually decreasing to zero by age 11·6. True conceptual thinking on the other hand is absent in the early stages but increases rapidly later until at age 11·6 it reaches 43 per cent of total responses and over 80 per cent of responses at the upper levels of the ability tests.

FIGURE 3

METHOD OF GROUPING, BY VERBAL TEST SCORE.



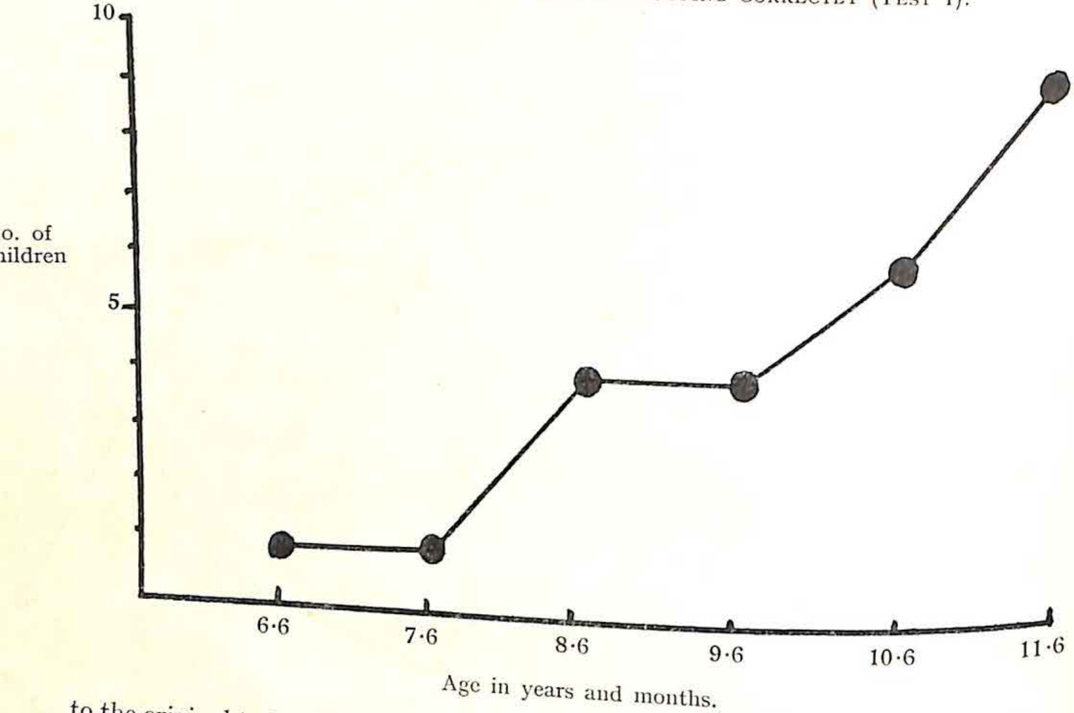
The Tests.

Figure 4 shows the number of subjects who regrouped the blocks correctly after concluding the original sorting. The improvement in performance with age confirms Vigotsky's findings. All except one of the subjects who showed any evidence of true conceptual grouping (phase three) were able to regroup the blocks successfully. Conversely, every subject except one who gave any evidence of operating at phase one (i.e., pre-complexive grouping) was unable to regroup successfully. Both the exceptional cases were marginal. The former regrouped successfully except for two blocks. As in one or two other cases this was almost certainly the accidental result of a too hasty regrouping since subjects who have achieved the concept often regroup the blocks very quickly and are prone to consequent errors in placing. The latter child made only one response at phase one (syncretic groupings) which again makes the case a marginal one.

Tests 1, 2 and 3 may be taken together as examples of transfer items. They are true extension tests since they demand the application of concepts to a problem involving discrimination among novel exemplars of the concept. The aggregate scores on these tests (104, 102, 100, maximum possible 240) were very similar, which suggests that the level of difficulty was much the same. Test one which demanded discrimination among three-dimensional models seemed to be at a very similar level of difficulty to tests 2 and 3 despite its closer similarity

FIGURE 4

NUMBER OF CHILDREN AT DIFFERENT AGES REGROUPING CORRECTLY (TEST 1).



to the original task. Test 4 which demands the *construction* of exemplars of the concepts was more difficult and produced only 69 (maximum possible 240) correct responses. Test 5, which demanded a verbal statement of the criterial qualities of the concepts symbolised by the nonsense syllables, proved most difficult and produced only 36 (maximum possible 240) correct answers. Intercorrelations among the tests were all high and signficant at the ·01 level.

In order to examine the relationship of the different levels of thinking to performance in the extension tests, correlation coefficients were computed between percentage of responses at the various levels and scores in the various tests. The resulting data are presented in Table 1. Here it can be seen that thinking at phases one and two correlates negatively with success in the extension tests, in most instances at a high level of significance.

TABLE 1

CORRELATION OF PROPORTION OF GROUPING AT THE THREE PHASES WITH PERFORMANCE IN THE EXTENSION TESTS.

	Tests					
	Regrouping	1	2	3	4	5
Phase 1	−0·36**	−0·25*	−0·33**	−0·36**	−0·29*	−0·21
Phase 2	−0·27*	−0·14	−0·16	−0·26*	−0·14	−0·42**
Phase 3	+0·38**	+0·35**	+0·40**	+0·50**	+0·34**	+0·58**

** Significant at ·01 level (one tail).
* Significant at ·05 level (one tail).

Grouping at phase three (true conceptual grouping) correlates positively and for the most part at a high level of significance with performance in the extension tests.

These data lend support to Vigotsky's findings concerning the stages of classificatory activity. They also suggest that the extension tests devised for the experiment were functioning satisfactorily.

IV.—DISCUSSION.

The principal aims of the enquiry were :

(1) to investigate the appropriateness of Vigotsky's model of concept formation with English primary school children ;

(2) to examine the relationships between levels of thinking and chronological age and mental age ;

(3) to investigate the relationships between level of thinking observed in the sorting experiment and performance in tests of extension.

With respect to (1) our findings bear out those of Vigotsky and other workers who have used the apparatus (Hanfmann and Kasanin 1937, Semeonoff and Trist 1958, Haslam 1966). Examples of the same three broad categories of grouping and most of the sub-categories were observed. However, we found little evidence of thinking at phase one, syncretic grouping. The negative correlation between operating at phase one and performance in the extension tests suggests very strongly that sorting at this level is not random. Random sorting at phase one would have produced a zero correlation with performance in extension tests. It would seem that children in this population operating in phase one, grouped according to contiguity or spatial organisation which operates counter to true classification, thus producing the negative correlation. The same can be said about phase two (complexive grouping). The correlations are somewhat lower than in phase one. This is almost certainly because at the highest level of complexive grouping the subject is operating in a similar way to phase three.

These correlational data indicate that it is possible for a child to complete a learning task with the help of corrections from the teacher and yet to have but a vague notion as to the principles involved in reaching a solution. The use of the apparatus and the extension tests in the experimental situation revealed the inadequacy of the child's performance. In many less structured learning situations this inadequacy would probably not be observed.

As compared with the findings of the Olver and Hornsby study (Bruner *et al.* 1966) the replacement of complexive by conceptual grouping occurs at a later age but this is presumably a function of the different material and the different approach. The general trend in the two studies, however, is very similar. In our study no fewer than ten children failed to group conceptually at all, and these were predominantly in the 6-year-old group. This result is similar to the findings of Annett and Olver and Hornsby who noted that children at this age tend to be satisfied with grouping in ' perceptible ' pairs, grouping on the basis of a single attribute, and are finished when one pair has been formed. This mode of grouping is more sophisticated than syncretic grouping, of which we found little with our sample. Haslam (1966) found with a very small sample (N=4) that all the moves made by children between 4 and 6 were of a syncretic nature. From her results it would seem that our sample comes into the band where complexive thinking predominates and that only the very youngest children are likely to give evidence of grouping of an entirely fortuitous nature.

In sum our conclusions to this section are that Vigotsky's approach to the investigation of the processes of concept formation is a fruitful one and that much the same modes of thinking which he observed seemed to be operating in English primary school children.

With respect to the second aim, our findings of a general trend towards predominantly conceptual grouping with increasing age, agree with the findings of other investigators. However, the trends illustrated by the relationship between the scores on the two intelligence scales are of particular interest. The trend observed when chronological age is considered, is made clearer when methods of classifying are related to score on Matrices but even clearer when related to Simplex scores (Figures 1 and 3). Neither of these scales is intended specifically to assess ability to classify but the fact that ability in performance with verbal material is more closely associated with sophisticated methods of classifying is interesting supporting evidence of Vigotsky's thesis of the importance of language in concept formation. Clearly, on this argument, proficiency in verbal skills should be closely associated with true conceptual thinking and this would be reflected in the actual manipulation of the concrete material. Our investigation seems to bear this out.

The results of the third prong of our investigation seem to us to be of particular interest. Vigotsky discusses the question of the extensive aspect of concept formation. He refers to the problems of applying a newly acquired concept to different exemplars and of giving a verbal definition of the concept. However, he gives no information about his tests of extension, presumably because he was more concerned with establishing the categories of thinking which were operating and their inter-relationship with language. Other clinical workers have been more concerned with using the apparatus as a test in itself. Since our interest was mainly pedagogical we were particularly interested in the way the subjects applied their newly acquired concepts. We considered that the extension tests would give us an indication of this. Our findings with respect to the extension tests do in fact show a close association between level of sophistication of grouping and ability to transfer.

One additional point which relates to all aspects of our study and to Vigotsky's views on the relationship between language and concept formation emerges from a scrutiny of Table 1. This table shows in a striking way how the ability to state the defining attributes of a concept is the acid test of conceptual thinking. Children who operate at phase two (complexive thinking) in the learning part of the experiment perform particularly badly when they have to verbalise the concept (test 5). Conversely, children who operate at phase three in the learning situation (conceptual thinking) are able to verbalise well and score highly on the extension tests. The non-significant correlation coefficient for phase one with test five is probably of less significance and suggests that responses in this test are closer to randomness because of the nature of the test which provides no clues of a visual nature at all. It would seem that this aspect of the experiment lends support to Vigotsky's argument that concepts are formed not merely through the interplay of associations "... but through an intellectual operation in which all the elementary mental functions participate in specific combination. This operation is guided by the use of words as the means of actively centring attention, of abstracting certain traits, synthesising them, and symbolising them by a sign."

The main conclusions of this study, therefore, demonstrate a close parallel between the classificatory behaviour and thinking in our sample and those found by Vigotsky. Further, we found that performance in a test of verbal ability was more closely related to level of sophistication of thinking than performance in a non-verbal test or chronological age, although all three were related. Finally, the ability to solve new problems making use of the same general principles which obtained in the concept formation task, was significantly correlated with the level of thinking demonstrated in the course of grappling with the task. It was also demonstrated that the ability to reach a 'correct' solution did not by itself guarantee that the subject was thinking conceptually according to the criteria demanded by the task.

V.—REFERENCES.

ANNETT, M. (1959). The classification of four common class concepts by children and adults. *Brit. J. Educ. Psychol.*, 29, 223-36.

BRUNER, J., *et al.* (1966). *Studies in Cognitive Growth.* New York: Wiley.

HANFMANN, E., and KASANIN, J. (1937). A method for the study of concept formation. *J. Psychol.*, 3, 521-540.

HASLAM, J. E. (1966). *A Study of Concept Formation in Children of Average Intelligence between the Ages of Six and Sixteen, Using Vigotsky's Method.* D.C.P. dissertation, Birmingham University.

INHELDER, B., and PIAGET, J. (1964). *The Early Growth of Logic in the Child.* London: Routledge.

PIAGET, J., and INHELDER, B. (1957). *The Growth of Logical Thinking.* London: Routledge.

SEMEONOFF, B., and LAIRD, A. J. (1952). The Vigotsky Test as a Measure of Intelligence. *Brit. J. Psychol.*, 43, 94-102.

SEMEONOFF, B., and TRIST, E. (1958). *Diagnostic Performance Tests.* London: Tavistock

VIGOTSKY, L. S. (1962). *Thought and Language.* New York: M.I.T. Press.

(*Manuscript received 29th February, 1968*)

OPINIONS ABOUT CAREERS AMONG SCOTTISH SECONDARY SCHOOL CHILDREN OF HIGH ABILITY

By H. J. BUTCHER and H. B. PONT

(*Department of Education, University of Manchester*)

SUMMARY. As part of a research into subject choice in Scottish senior secondary schools, 1,100 pupils were required to rate 15 careers on 6 criteria and to write essays about their future careers. Boys and girls agreed quite closely on, for example, the usefulness and prestige of the careers, but differed markedly about liking and interest. Boys put 'engineer' highest on 'liking,' but much lower on 'salary' and prestige; girls put 'teacher' (of non-science subjects) highest on 'liking,' but much lower on salary and prestige. Similar results emerged from analysis of the essays.

I.—INTRODUCTION.

THIS study forms part of a larger research into choice of specialisation in secondary school and university and into the causes of a relative shortage of entrants to university science departments. It will not be primarily concerned with these broader issues, but will summarise data from two of the many kinds of information collected—from ratings by school children of 15 specified careers in terms of six criteria, and from essays written about what they expected to be doing in five years' time.

The group of children studied in the main research numbered about 1,100 in twenty schools (including local authority, grant-aided and independent schools) and was intended to represent the population of children in Scottish schools likely eventually to become students at Edinburgh and Heriot-Watt Universities. The twenty schools were chosen from those that had sent most students to the two universities in recent years. Within these schools, about half the age-group was selected, so as to include the pupils thought most likely to reach a university. The testing and collecting of information took place during the session 1965-66, when all the children were in their second year in secondary school.

The results reported here are based not on the entire group of 1,100, but on random samples, stratified by school, of 150 boys and 150 girls.

II.—ADMINISTRATION OF THE RATING SCALES.

The fifteen careers about which opinions were obtained were: Lawyer, Engineer, Doctor, Journalist, Research Chemist, Minister of Religion, Biologist, Physicist, Psychologist, Personnel Manager, Insurance Executive, Science Teacher, Social Worker, Banker, Teacher (of non-science subjects).

The subjects were required to rate these in six different ways, the six criteria being: (1) how much they would like such a career (P); (2) how interesting the career would be (Q); (3) how high a salary would be earned (R); (4) how much prestige the career would bestow (S); (5) how high would be the qualifications needed (T); and (6) how useful the career would be to society (U).

Each rating was on a five-point scale, and verbal indications were provided to show the direction and meaning of the ratings, e.g., "4 means 'I would like it very much'." Pupils rated all careers on one criterion before proceeding to the next.

The task was a heavy one, requiring 90 ratings from each child, and it might be supposed that too many discriminations were demanded. All the evidence suggests, however, that the work was carried out carefully, sensibly and with a good deal of interest. The resulting judgments appear reliable and meaningful.

III.—RESULTS.

Table 1 gives the average ratings by boys and girls separately of the 15 careers on the six criteria. The rank position of each career on a criterion is given in brackets after the mean.

TABLE 1

RATINGS OF 15 CAREERS BY 300 SCOTTISH CHILDREN.

Mean ratings on five-point scale (0—low, 4—high) and rank order of preference.

Occupations		Like (P)		Interest (Q)		Salary (R)		Prestige (S)		Qualif. (T)		Useful. (U)	
		M	Rank	M	Rank	M	Rank	M	Rank	M	Rank	M	Rank
Lawyer	B	1·87	(10)	2·39	(8)	3·33	(1)	3·30	(2)	3·59	(3)	3·25	(3)
	G	1·68	(10)	2·45	(8)	3·50	(1)	3·44	(2)	3·68	(3)	3·25	(4)
Engineer	B	2·50	(1)	2·79	(3)	2·79	(8)	2·60	(7)	3·01	(9)	3·16	(4)
	G	1·06	(14)	1·73	(13)	2·63	(12)	2·36	(14)	2·82	(9)	2·69	(11)
Doctor	B	2·21	(3)	2·93	(2)	2·99	(6)	3·41	(1)	3·85	(1)	3·91	(1)
	G	2·48	(2)	3 14	(1)	3·14	(4)	3·63	(1)	3·92	(1)	3·84	(1)
Journalist	B	1·93	(8)	2·64	(5)	2·48	(12)	2·16	(15)	2·19	(14)	2·31	(14)
	G	2·38	(3)	3·01	(3)	2·50	(13)	2·31	(15)	2·36	(14)	2·29	(14)
Research Chemist	B	2·45	(2)	3·05	(1)	3·18	(2)	3·04	(3)	3·72	(2)	3·44	(2)
	G	1·91	(8)	2·69	(5)	3·19	(2)	3·11	(3)	3·71	(2)	3·29	(2)
Minister	B	·95	(15)	1·17	(15)	2·03	(14)	2·57	(9)	2·51	(11)	2·65	(11)
	G	1·01	(15)	1·48	(14)	2·08	(14)	2·88	(6)	2·79	(10)	2·94	(8)
Biologist	B	2·04	(6)	2·67	(4)	2·76	(9)	2·58	(8)	3·40	(6)	2·96	(7)
	G	1·93	(7)	2·48	(7)	2·85	(10)	2·69	(10)	3·35	(7)	2'83	(10)
Science Teacher	B	1·96	(7)	2·22	(10)	2·49	(11)	2·42	(11)	3·25	(7)	2·83	(9)
	G	1·42	(13)	1·84	(11)	2·95	(7)	2·70	(9)	3·47	(4)	2·87	(9)
Social Worker	B	1·47	(14)	1·76	(12)	1·70	(15)	2·26	(14)	1·95	(15)	2·80	(10)
	G	2·36	(4)	2·51	(6)	1·94	(15)	2·49	(11)	2·09	(15)	3·04	(6)
Banker	B	1·87	(11)	1·66	(13)	3·00	(5)	2·73	(6)	2·70	(10)	2·58	(12)
	G	1·86	(9)	1·77	(12)	3·05	(5)	2·77	(8)	2·62	(11)	2·59	(12)
Physicist	B	2·18	(4)	2·53	(6)	3·06	(4)	2·93	(4)	3·50	(4)	3·07	(5)
	G	1·55	(11)	1·97	(10)	2·96	(6)	2·87	(7)	3·34	(8)	2·92	(7)
Psychologist	B	1·84	(12)	2·51	(7)	3·09	(3)	2·79	(5)	3·34	(5)	2·92	(8)
	G	2·16	(5)	2·74	(4)	3·23	(3)	3·01	(4)	3·45	(5)	3·14	(5)
Personnel Manager	B	1·93	(9)	1·88	(11)	2·90	(7)	2·42	(10)	2·47	(13)	2·31	(13)
	G	1·98	(6)	2·09	(9)	2·85	(11)	2·51	(11)	2·43	(13)	2·36	(13)
Insurance Executive	B	1·57	(13)	1·61	(14)	2·72	(10)	2·33	(13)	2·49	(12)	2·15	(15)
	G	1·47	(12)	1·42	(15)	2·87	(9)	2·43	(13)	2·57	(12)	2·27	(15)
Teacher (OTS)	B	2·13	(5)	2·33	(9)	2·36	(13)	2·34	(12)	3·07	(8)	3·01	(6)
	G	2·90	(1)	3·05	(2)	2·92	(8)	2·99	(5)	3·40	(6)	3·28	(3)

An obvious feature of these figures is the general similarity, on criteria R, S, T and U, between the ratings of girls and boys. The differences that stand out most clearly are that boys rate higher on ' liking ' and ' interest ' the careers of engineer, research chemist, physicist and science teacher (all concerned with science), whereas girls rate more highly those of social worker, teacher of non-science subjects, journalist and psychologist (all interpretable as more concerned with people).

The sex differences are much more evident in the ratings on the first two criteria, those concerned with liking and interest, than on the last four, which call for more objective assessments. This seems to indicate that the judgments have been carefully considered.

The ordering of the careers will be discussed in a later section, but certain general tendencies are immediately clear without detailed analysis. 'Lawyer' and 'doctor' are rated high on all criteria, the latter being put almost unanimously far above other professional men in usefulness to society. 'Minister of Religion' is put by far the lowest in terms of liking, although fairly high on, for instance, usefulness to society. The 'business' or less academic careers (banker, insurance executive, personnel manager) although quite well liked, are not placed very high on prestige or usefulness to society. A career in journalism is thought of as decidedly interesting, but as very low in prestige and usefulness.

Overlap between the criteria.

Most of these interpretations assume that the ratings on the six criteria represent genuinely different bases of judgment and are not the same thing under six different names. Correlations were therefore calculated of each criterion with every other, averaged over the 15 careers. This was done for boys and girls separately, but the two patterns were closely similar, and the figures in Table 2 below are for the combined group.

TABLE 2

INTER-CORRELATIONS BETWEEN THE RATING CRITERIA
(Boys and Girls. N=300)

	P	Q	R	S	T	U
P	1·00	·71	·26	·30	·25	·29
Q		1·00	·28	·34	·34	·35
R			1·00	·45	·50	·25
S				1·00	·49	·46
T					1·00	·49
U						1·00
	P—Like. S—Prestige.		Q—Interest. T—Qualifs.		R—Salary. U—Usefulness.	

It can be seen that in Table 2 the first two criteria correlate highly, but that correlations between the first two and the last four and also within the last four are markedly lower.

This makes good sense. The first two criteria ('liking' and 'interest') are intended to assess personal taste and inclination, but the remaining four are meant to be more objective assessments of the nature and status of the various careers. From Table 2 one can conclude that not very much is to be gained by distinguishing ratings on the first two criteria, but that the last four seem genuinely to represent different bases of judgment and not to be vitiated by 'halo effect.'

A multiple regression analysis was carried out to determine how far the first and most general rating, that of 'liking' could be predicted from the other five, and with what weights. As one would suppose from inspection of the correlation coefficients in Table 2, 'interest' was greatly predominant as a predictor of 'liking.' The multiple R, at 0·72, was hardly larger than the correlation with 'interest' alone. It was then thought that a clearer picture of the relative importance of the last four bases of rating might be obtained by removing 'interest' and its apparent masking effect from the analysis. The results are shown in Table 3; they give a multiple R of 0·369.

TABLE 3

'LIKING' FOR A CAREER AS PREDICTED BY THE FOUR MORE 'OBJECTIVE' RATINGS.

	β_{Pk}	r_{Pk}	$\beta_{Pk} r_{Pk}$
R	·1424	·2610	·0372
S	·1385	·2971	·0411
T	·0223	·2504	·0056
U	·1811	·2915	·0528
	R—Salary. T—Qualifications.	S—Prestige. U—Usefulness.	

From Table 3 it appears that the relative contributions of the four more 'objective' assessments to overall 'liking' of a career, whether measured by beta weights or by products of beta weights and correlations, are as follows: None is very high, but the descending order of importance is 'usefulness to society,' then 'salary' and 'prestige' about equal, and finally (a very small contribution, as seems reasonable) 'qualifications required.'

The rank ordering of careers.

The ranking on any particular criterion is shown in brackets for boys and girls separately in Table 1. Besides the rank order, means of ratings are given. Inspection of the means will show where the rank order alone may be misleading, as where two careers receive an almost identical rating; the standard deviations are omitted for reasons of space, but the variations observed in spread of ratings deserve a brief comment.

To take the latter point first, it is clear that there is far more agreement about some careers (on some criteria) than about others. The highest agreement, both among boys and girls, is about the usefulness to society of doctors. The greatest disagreement among both is about liking for the career of research chemist—presumably because this is particularly attractive to many and particularly unattractive to many others.

More generally, one can see that in both tables the level of agreement is noticeably higher, as one would hope, in the four more 'objective' ratings than in the first two, i.e., 'liking' and 'interest' This is shown in Table 4 below, where the standard deviations of the ratings have been averaged over careers.

TABLE 4

STANDARD DEVIATIONS OF RATINGS AVERAGED OVER CAREERS.

	P	Q	R	S	T	U
Boys	1·16	1·14	0·76	0·92	0·72	0·87
Girls	1·02	1·03	0·71	0·77	·69	0·82
Total ..	1·09	1·09	0·73	0·84	0·70	0·84
	P—Like. S—Prestige.		Q—Interest. T—Qualifications.		R—Salary. U—Usefulness.	

Another tendency that emerges, and one that would hardly have been anticipated, in the consistently higher agreement in rating among girls than among boys.

More important, however, is the rank ordering of careers. Bearing in mind the aims of the larger research within which this study of opinions was carried out, we shall pay particular attention to the rating of careers in science.

In Table 1 it is at once noticeable that in the first column, where careers are ranked on average 'liking' by boys, five science careers are among the first seven. Similarly, on the basis of 'interest,' five are among the first eight. This is to claim that 'psychologist' represents a science career, which is perhaps justifiable in the present context; the profile is more like those of chemist and physicist than like that of social worker. In columns 3 and 4, however, showing rankings on salary and prestige, the position is slightly different. Only three and four scientific careers appear respectively among the first seven. This contrast is much more marked if one considers the career of engineer. It is the most liked of all those listed, but is not placed in the first six in ratings on salary or prestige. The realism of these children's views is confirmed by the recent publication of the Jones report (1967) in which it is noted that the emigration of engineers and technologists to North America has quadrupled during the last six years (and this is ascribed directly to differences in salary and prestige).

The girls place only two scientific careers among the seven liked best; psychologist is in fifth place and biologist in seventh. In their ratings on salary and prestige there is greater agreement with the boys. Again three or four appear among the first seven.

Two main conclusions therefore emerge fairly clearly about the rating of careers in science compared with other careers. Firstly—and this is far from being a new finding, but is here very strongly confirmed—there is a predominant sex difference in liking for scientific careers, very evident at age 13. Secondly, among boys, from whom the vast majority of graduate scientists will be drawn, there is a strong inclination towards a scientific career, including particularly a career in engineering, but already a realistic assessment and expectation of rather low salary and prestige.

Preferences among the scientific careers can be seen more clearly in Tables 5 and 6. Table 5 seems to confirm the widely held stereotype of the engineer in so far as he is thought of as less well qualified than the other scientists. In this table and also in Table 6 (girls' preferences) the career of research chemist is rated very highly. Possibly the magic word 'research' has accounted for some of this acclaim, although it was added to avoid possible confusion between 'chemist' and 'pharmacist.' It is slightly surprising that the career of 'physicist' is rated below that of 'biologist' for interest by the boys; it is possible that this reflects a changing climate of opinion and an increasing interest in the biological sciences.

TABLE 5

SCIENTIFIC CAREERS. ORDER OF RATING BY BOYS. N=150.

	P	Q	R	S	T	U
Engineer	1	2	3	3	5	2
Research Chemist..	2	1	1	1	1	1
Physicist	3	4	2	2	2	3
Biologist	4	3	4	4	3	4
Science Teacher ..	5	5	5	5	4	5

P—Like. Q—Interest. R—Salary.
S—Prestigage. T—Qualifications. U—Usefulness.

TABLE 6

SCIENTIFIC CAREERS. ORDER OF RATING BY GIRLS. N=150.

	P	Q	R	S	T	U
Biologist	1	2	4	4	3	4
Research Chemist..	2	1	1	1	1	1
Physicist	3	3	2	2	4	2
Science Teacher ...	4	4	3	3	2	3
Engineer	5	5	5	5	5	5

P—Like. Q—Interest. R—Salary.
S—Prestige. T—Qualifications. U—Usefulness.

Science teaching is rated low as a career by both boys and girls. For scientifically inclined boys it is apparently the least popular choice among the five scientific careers listed. Among scientifically inclined girls, only engineering is less liked. The relative distaste of girls for science teaching is the more marked when one remembers that teaching of other subjects was the career best liked by girls of all the 15 (Table 1).

Career choices reported in the essays.

All the results reported so far have been obtained from the pupils' ratings of a limited list of careers. This section describes results obtained when they provided an open-ended account of their occupational aspirations by writing an essay entitled " What I hope to be doing in five or six years' time." It was possible to analyse these essays in two ways. Firstly the frequency of reporting, either as a first choice or as an alternative, of all careers was noted, and secondly, the reasons for the first choice, either stated or implied, were tabulated. Whereas the analysis of the ratings was based on a random sample of 300 out of 1,100+ children, the following account deals with the whole group, or rather with the 937 from whom essays were received (450 boys, 487 girls). The career choices mentioned five or more times are given separately for boys and girls in Table 7. 49 boys and 51 girls (about 11 and 10 per cent respectively) indicated in their essays that they did not know what they wanted to be.

The proportion of science choices reported is very relevant to the aims of the whole research. Although the main branches of science are among the careers most often mentioned the total percentage of science choices is, for boys, about 38 per cent. Among girls the only science careers in the 15 most often mentioned are ' research scientist ' and ' science teacher ' and the overall percentage of science careers chosen is about 12. A difference in the academic level of career choice between boys and girls is apparent. Among the boys about 88 per cent. indicated a choice of career requiring university or equivalent qualification, but among the girls only 66 per cent and of these 22 per cent indicated a preference for primary teaching.

In about one half of the boys' essays and about one third of those of the girls no reasons for the choice were given. In the remainder it was found that the reasons given reflected largely the findings of the rating scale in that an interest in the subject involved appeared as the deciding factor, while reasons such as salary, etc., were of much less importance.

TABLE 7

CAREER CHOICES.

Boys (N=450)		Girls (N=487)	
Career	No. of times reported	Career	No. of times reported
Engineer	66	Secondary Teacher (not Science)	112
Teacher (not Science)	45	Primary Teacher	109
Research Chemist	33	Secretary	43
Doctor	31	Research Scientist	28
Architect	27	Nurse	27
Physicist	24	Civil Service	24
Lawyer	21	Science Teacher	22
Scientist	19	Air Hostess	20
Pilot	19	Veterinary Surgeon	19
Science Teacher	16	Social Worker	17
Services	16	Child Nurse	15
Veterinary Surgeon	15	Interpreter	14
Draughtsman	13	Doctor	14
Civil Service	13	Journalist	13
Journalist	13	Dress Designer	9
Farmer	12	Laboratory Technician or Assistant	6
Chartered Accountant	12	Radiographer	6
Electrician	9	Physiotherapist	5
Minister	8		
Biochemist	6		
Police	6		
Chef	5		
Interpreter	5		
Banker	5		
Footballer	5		
Biologist	5		
Dentist	5		

IV.—CONCLUSION.

A few words are required to put this study in its context. The Dainton Committee report (1968) finds the Scottish system of secondary and higher education markedly more satisfactory than the English in its avoidance of premature specialisation. It sounds a note of warning, however, in its comment that the introduction of the Ordinary grade examination resulted in some degree of specialisation one year earlier than had previously been the case; also in saying (sect. 93) that " we strongly hope that the introduction of this new Certificate (of sixth form studies) will not lead to pressure away from the traditional broad base of the secondary school course in Scotland."

The present paper has shown that among the most able second-year secondary school pupils there is considerable interest in and liking for careers in science (among boys at least). From other preliminary findings in the research of which this forms a small part, we suspect that even in the comparatively flexible Scottish system there may be some wastage of potentially good scientists (arising principally from the entrance requirements of science faculties in Scottish universities and the effects of these requirements on subject choice at the beginning of the third year in secondary school). Some of these effects are described by Pont and Butcher (1968), but the existence and extent of such possible wastage cannot be confirmed until our data have been more fully analysed.

ACKNOWLEDGEMENT.—This research was supported financially by the Scottish Council for Research in Education. The writers wish also to acknowledge help and encouragement from its Director, Dr. D. A. Walker.

V.—REFERENCES.

DAINTON REPORT (1968). *Enquiry into the Flow of Candidates in Science and Technology into Higher Education*. Cmnd. 3541. London : H.M.S.O.

JONES REPORT (1967). *The Brain Drain*. Cmnd. 3417. London : H.M.S.O.

PONT, H. B., and BUTCHER, H. J. (1968). Choice of course and subject specialisation in seventeen Scottish schools. *Scottish Educational Studies*, 1 (2), 9-15.

(*Manuscript received 19th February*, 1968)

THE IMAGE OF THE TEACHER

By JULIA H. CLARKE (J. H. Pheasant)
(*University of London Institute of Education*)

Summary. It has been suggested that if salaries paid to teachers were competitive with those offered by industry, the present shortage of certain categories of teachers, such as teachers of science and of handicraft and infant teachers, would not exist. There is evidence from young people considering the choice of a career that the financial status of a teacher is only one factor deterring them from training to teach. Other factors which seem to deter both boys and girls, such as their fear of undisciplined classes and monotonous work are discussed. Ways in which adolescents and their parents may become better informed regarding the life of a teacher and teacher education are suggested.

I.—Introduction.

The shortage of teachers of science, mathematics and handicrafts in our secondary schools and of qualified infant teachers is sometimes considered to be a state of affairs which could be overcome simply by offering teachers higher salaries. Some of those who seek to solve this educational problem believe, moreover, that if the country cannot afford to pay all teachers highly, the methods of industry might be followed by which those categories of teachers who happen to be in short supply would be paid more highly than their colleagues. The shortage of science teachers would not exist, it is argued, if salaries offered to intending teachers were large enough to be competitive with those offered by industry and young women would welcome the opportunity to become infant teachers. It seemed desirable to find evidence for or against this solution to the problem.

II.—Evidence from young people in their pre-sixth form year.

It is reasonable to suppose that information concerning the grounds on which young people select or reject teaching as a career might be relevant and could be obtained from those secondary school pupils expected by their headmasters to attain the necessary examination qualifications for entry into a university or training college. A certain amount of evidence was collected recently during investigation into the influence of school organisation and curriculum on the choice of examination subjects and careers, involving 1,511 pre-sixth form pupils (634 girls and 877 boys) from 18 secondary schools constituting a small but reasonably representative sample of schools submitting pupils for both O and A levels of the General Certificate of Education (Pheasant 1960).

Of these pupils, mainly between the ages of 16 and 17 years, 110 boys and 33 girls had come to no decision about their future employment, but of the rest (i.e., 1,368 pupils) only 35 boys and 115 girls stated that they hoped to teach. Only eight of the boys and 11 of the girls wished to teach science or mathematics and 27 boys and 50 girls hoped to teach other subjects in secondary schools. 54 girls but no boys wished to teach in primary schools. 23 of the boys and 36 of the girls expected to obtain a degree before teaching ; while 12 boys and 79 girls planned to enter training college and to teach as non-graduates. Thus, of all the pupils who had indicated their choice of a career, 21·7 per cent of the girls and 5·2 per cent of the boys wished to teach. Of those who chose careers involving science or mathematics, 11·7 per cent of the girls and only 1·9 per cent of the boys were willing to teach.

It was clear from the evidence provided by questionnaires and interviews that most of the boys and girls looked for a career in which they could use a school subject in which they had become interested. Some boys and many more girls wished to work with and for people. It seems surprising that, since teaching would appear to offer excellent opportunity for work with people and for the use of a favourite academic subject, the majority of pupils who stressed these two conditions chose careers other than teaching. The careers chosen by such pupils are listed in Table 1. In order to compare the frequencies of girls and boys, the number of girls choosing each sphere of employment has been multiplied by 1·26.

TABLE 1

CAREERS CHOSEN BY PUPILS.

Employment calling for interest in people and use of an academic subject	Chosen by		Girls ×1·26
	Boys	Girls	
Teaching	35	115	145
Clerical work	30	147	185
Social work	6	20	25
Modern Language and Travel	11	34	43
Buying and Selling	19	9	11
Civil Service and Administration	24	15	19
Nursing and Medical Auxiliaries	0	65	82
Medical Sciences	67	28	35
N (Total N in brackets)	192 (877)	433 (634)	545 (877)

Desire to work with and for people seemed of over-riding importance to the girls, most of whom stated that satisfying work was more important to them than a high salary. They were interested in Arts subjects, especially modern languages rather than in Science: perhaps an expression of their interest in people. The largest group was seeking secretarial posts after one year's training. Early marriage was expected and the long training needed for nursing or teaching was not considered practicable. Office work was thought to offer congenial employment and company; to provide "enough money for buying clothes, helping mother a bit financially before I leave home and to save something towards getting married." Encourage to look beyond the immediate future, many girls stated that they thought they would enjoy being at a teachers' training college; they knew that grants were available for training; they appreciated the advantages to a working mother of having employment which involved understanding children and which enabled her to be free during her children's holidays. Nevertheless, the majority indicated that they would not like teaching but there seemed considerable reluctance to say why this was so.

The usual first reaction of the boys interviewed to the question why they had rejected teaching was that "the salary is not high enough" or "because I can get a much better salary in industry or commerce." When asked if they would teach if salaries were comparable with those in industry, all these boys replied that they would not. When asked to give their reasons the most frequent reply was that chances of promotion in schools were restricted and unattractive. When pressed further, they all said that they would not teach even if prospects in the teaching profession were as good as they expected them to be in industry. It would appear that whereas the length of training needed for teaching was a drawback to girls and poor prospects of salary and promotion were hindering boys from taking up teaching there were also other factors which pupils were unwilling to mention.

III.—Evidence from sixth formers.

More information as to why young people were rejecting teaching was collected from pupils in the second year of their sixth form course. Pupils in the science sixth form of six schools not included in the eighteen used already were asked to state anonymously on paper their sex, their proposed careers and their reasons for accepting or rejecting the idea of taking up teaching. 14 of the 29 girls hoped to teach in junior schools. They rejected infant school work because they thought that they could not use their science. Eight of them had considered teaching science in secondary schools but feared discipline problems. The remaining 15 girls mentioned that they did not wish to undertake long training and that they feared they could not keep order if they tried to teach. They were intending to become laboratory technicians after taking their A level examination. Only six of the 175 boys proposed to teach. In addition to the previously listed reasons for rejecting teaching (the low salary of teachers and poor prospects of promotion), these sixth formers mentioned the monotony of teaching, the fear of not controlling a class and dissuasion by their parents or teachers. One boy wrote " The status of a graduate in school is lower than in industry."

IV.—The image of the teacher.

Each boy and girl making a career choice had been in close proximity to a number of teachers during his or her school life. From this experience the adolescent had built up a picture of the teacher's life. Some typical comments from the sixth formers were " Year after year he teaches troublesome children the same things for next to nothing in wages " and " Trying to teach dim, rebellious children does not appeal to me. . " It is all right if a teacher can keep order but boys despise masters who cannot." It would seem that the fear of unruly classes is a factor which may be deterring young people from teaching. The younger pupils had indicated that certain teachers whose lessons they disliked could not control classes, whereas others whose lessons were most stimulating had no difficulty with discipline. The image of a teacher tormented by pupils and unable to make satisfactory contact with them, is reinforced by the idea of his dull, repetitive, monotonous job continuing year after year with little prospect of promotion to a post carrying a salary which might approximate to that offered to successful men in industry. The adolescent (whose dislike of subjects which bore him was expressed freely during the work with pre-sixth formers) rejects the idea of teaching. He looks instead for other work in which he can be with adults and can use a subject he enjoys in a job which offers promotion in proportion to his value to his employers.

Boys and girls might consider teaching if more were done to dispel some of their misconceptions concerning the teacher's life and the relative glamour of competing careers. There is much dreary routine and repetitive work in industry and in the professions which seem more attractive to pupils. More publicity might be given to the experience of those young men and women with science degrees who, after going to industry, subsequently left to become teachers.

V.—Evidence from near-graduates.

Unprompted reasons for choosing teaching as a career were obtained from 96 university science students in the last year of their degree course who had applied for admittance to a post-graduate certificate course of teacher training. The responses were generally similar for both sexes. 95 expressed a wish to work with people; 87 stated that they had a wide interest in science rather than an interest in a small field of science; 68 said that they had enjoyed learning and wanted to help others to do so; 47 liked "explaining, getting things across"; 81 said that they had enjoyed their own school days; and 15 (all women) said, "Teaching is a job you can come back to when your children start school."

72 said that teaching now had the added attraction that part-time research under the auspices of the universities and the British Association was possible for them in school laboratories. This they felt would "keep them out of a rut": like the adolescents, they feared monotony. Remarks relevant to their choice of teaching were: "Masters at school told me not to teach but the staff at college think highly of teaching and recommend it." "My parents are teachers. I have been in and out of their schools since I was in the primary school. I have always wanted to teach." "I like being with people and now that the universities are helping with part-time research I feel that I can both teach and do some research as long as I have a reasonably good laboratory at school." "The science masters at school tried to put me off teaching; they said that if you had a good degree you were wasting your time for industry was better paid and less monotonous."

VI.—EVIDENCE FROM HEADS OF SCHOOLS AND TEACHING STAFF.

Discussion with heads of schools and with teachers showed that most of those interviewed enjoyed teaching and were surprised when told that pupils considered that teaching was monotonous and that teachers were beset by fears of undisciplined classes. One teacher whose subject attracted more sixth formers than did any other in that school said "Do they really see us like that?" Nevertheless, though these teachers enjoyed their work they were, on the whole, reluctant to recommend their pupils to take up teaching. One head remarked, "I enjoy teaching; I enjoy running a school. But there is so much uneasiness among teachers to day: so much dissatisfaction over prospects, over teaching large classes, over lack of co-operation from parents of those pupils who do not fit in with the life of the school. Many teachers are disturbed by the possible effects of the reorganisation of schools for what they believe to be political rather than educational reasons. While this situation exists, one cannot wholeheartedly encourage bright young people to join the teaching profession."

VII.—ENJOYMENT OF SCHOOL.

That happiness at school might influence pupils to choose teaching as a career was suggested by remarks made by girls in a school used in the pre-sixth form sample. Of the 98 girls of this school who were involved, 25 hoped to teach. Of these, 18 stated that they had first considered teaching while they were primary school pupils and their tentative choice at an early age had been confirmed by their experience of life in their secondary school. The headmistress of this school reported that the 40 primary schools in the area were for the most part happy schools. This headmistress is a serene person; her staff contains a high proportion of lively teachers including a considerable number (around 25 per cent) of part-time married women who have been welcomed to the school on their return to teaching and who contribute fully to the school's activities both in the classroom and in school societies. A bright, attractive, though overcrowded school building may also contribute to a situation which seems to influence girls to take up teaching.

In another school, teaching was the career chosen most frequently by the girls of the sample (15 out of 58) and ranked second in popularity among the boys (6 out of 34), the majority of whom chose some form of engineering. It might be argued that in the rural hinterland of this Welsh school there was less variety of available employment than in a large conurbation, or that the parents of these pupils valued education more highly than did the parents of pupils attending the same type of school in urban areas. However, it is interesting to note that only two out of the 92 pupils from this school stated that they were bored by any school subject ; none expressed a dislike of school and all but one stated that they enjoyed school. On the other hand, many pupils from other schools of the sample stated that they were bored with some subject or subjects and enjoyment of school was rarely mentioned.

VIII.—SOURCES OF HELP WITH CAREER CHOICE.

Some pre-sixth form pupils stated that the usual careers lectures given to fifth formers were too general to be of much assistance in helping them to select a career. They said that they would have preferred a series of informal talks, each about one or two related careers. Teaching as a career had been mentioned briefly in the lectures given by visiting careers advisers and it had been suggested that the school staff would give information about training. It was assumed by the advisers that the work of a teacher was familiar to adolescents whereas information about training, conditions of work and prospects offered by other forms of employment were not known by school leavers and needed to be brought to the attention of young people and their parents.

Each of the pre-sixth form pupils was asked to select from a list of fourteen sources, those which had aided him most in choosing a career. If a pupil acknowledged help from two or three sources he was asked to rank such help in order of importance For scoring the responses, three points were allotted to each first choice, two to a second and one to a third choice. Responses were obtained from 795 boys and 560 girls. Of these, 94 boys and 32 girls said that they had not yet chosen a career. The scores therefore represent the ratings made by 701 boys and 528 girls.

TABLE 2

ACKNOWLEDGED SOURCES OF HELP IN CHOOSING A CAREER.

Sources of Help	Scores (First choice three points)		
	Boys	Girls	Total
Father	948	478	1426
Mother	509	595	1104
Other members of the family and family friends	400	225	625
School friends	62	49	111
School Staff	305	295	600
Y.E.O., Minister of Religion, Careers lectures, literature about careers	534	369	903
Number acknowledging help	701	528	1229
Unaided choice, or no significant help acknowledged	144	46	190

In the sample as a whole, parental influence on the choice of a career appeared to outweigh all other sources of help. It was greatest in the case of pupils from independent and direct grant schools where it accounted for 58 to 71 per cent of the sources acknowledged, while in the maintained schools of the sample the range was from 33·9 to 55·4 per cent, It was evident, however, that many parents of pupils interviewed knew little of life within the school building attended by their sons or daughters. In one school the headmistress sought to overcome this ignorance of school life by arranging frequent sessions at which parents could meet their daughters' teachers. Parents were helped to understand the aims of the school, the methods of teaching used and the modern content of the curriculum subjects. They were invited to join in school functions and they used the school swimming pool which they had helped to build and equip. They were being given the opportunity to co-operate with the teachers as partners in their daughters' education.

IX.—Discussion.

It would appear that many adolescents are rejecting teaching as a career partly because of their fear that the work will be monotonous and that they might be unable to maintain order and respect in the classroom. Careers advisers have tended to assume that any young person thinking about the possibility of becoming a teacher would obtain information and answers to his questions from his own teachers. Moreover, parents who appear to influence the choice of their children's careers to a considerable degree, frequently have little or no contact with the schools and so would be unlikely to draw the attention of young people to a teaching career. Some of the teachers expressed the view that they found difficulty in discussing the training and work of a teacher with their own pupils but said they would welcome an informal session held in the school, when adolescents and their parents could meet and discuss with teachers the questions raised by young people wondering whether or not to enter teaching.

It is hoped that the introduction of new syllabuses, of new methods of teaching, of pupil-participation in such activities as project work and of new methods of examining, all of which are currently engaging the attention of teachers, may result in more acceptable learning situations for the pupils and this in turn may lead to an increase in the proportion of young people who are prepared to teach.

X.—References.

Collins, M. (1964). Untrained and trained graduate teachers. *Brit. J. Educ. Psychol.*, 34, 75-84.

Evans, K. M. (1952). A study of attitude towards teaching as a career. *Brit. J. Educ. Psychol.*, 22, 63-69.

Floud, J. E., Halsey, A. H., and Martin, F. M. (1956). *Social Class and Educational Opportunity*. London : Heinemann.

Gallagher, M. (1957). The prospect of marriage. *Brit. J. Educ. Psychol.*, 27, 248-24.

Himmelweit, H. (1954). Social status in secondary education. In Glass, D. V. (ed.), *Social Mobility in Britain*. London : Routledge.

Pheasant, J. H. (1960). *The Influence of School Organisation and Curriculum on the Choice of Examination Subjects and Careers*. Ph.D. thesis, Univ. of London.

Pheasant, J. H. (1961). The influence of the school on the choice of science careers. *Brit. J. Educ. Psychol.*, 31, 38-42.

Porter, J. F. (1957). *A Sociological Study of the Causes of Early Leaving of Children from Grammar Schools*. M.A. Thesis, Univ. of London.

(*Manuscript received* 17*th May*, 1966)

PUBERTY, FAMILY SIZE AND THE EDUCATIONAL PROGRESS OF GIRLS

By PAMELA K. POPPLETON
(*Department of Education, University of Sheffield*)

SUMMARY. The aim of this study was to examine the influence of two variables: rate of physical maturation and family size on the intelligence and attainments of girls between the ages of 11½ and 15½ attending schools in the West Riding of Yorkshire. Differences in performance on Test AH4 and attainment in English and Mathematics were shown to exist between groups classified according to age at puberty but most differences were not statistically significant and could be accounted for by the fact that early maturing girls came from smaller families than late maturing girls. Correlations were calculated between the educational measures and four social and developmental variables and the results submitted to factor analysis. Two factors were identified: a social-educational factor and a developmental factor. Family size contributed to both factors but contributed more to the educational variance than did age at puberty. It is suggested that rate of physical maturation and family size are related and act jointly to produce the effects formerly attributed to age at puberty alone.

I.—INTRODUCTION

THE relationship of physical and mental growth is a continuing source of interest to all who are concerned with the study of child development It comes to be a subject of great importance at adolescence when, it has been suggested, early maturing children may have considerable advantages over their later maturing age-mates both educationally and psychologically. Tanner (1961) for example, suggested that "a fast maturing boy will have a somewhat greater chance of passing any age-bound examination than will a slow-maturing boy of the same chronological age . . . To be fair, either some estimate of developmental age should be made and this factor allowed for, or else, more practically, a series of opportunities for the later maturer to catch up should be provided." This argument was particularly relevant to the 11 plus examination and, though this hurdle is now fast disappearing from English education, it still applies to age of transfer decisions (Plowden Report, 1966). But does the late maturer catch up, and is it simply a question of providing more opportunities? If so, the extension of the school leaving age would be important for late maturing children. The problem is essentially whether there is an adolescent spurt in mental growth which parallels the spurt in physical growth just before and during adolescence; and the evidence relating to this problem has been exhaustively examined by Nisbet and Entwistle (1966). They conclude that the differences which exist at puberty can be traced back to the age of 7 and that any spurt is small and insignificant. Douglas (1964) came to the same conclusion and the situation now appears to be that while the rate of physical maturation is related to educational progress, groups of early, average and late maturing pupils maintain their rank order throughout the period of formal schooling. On the basis of the National Survey findings however, Douglas adds a qualification: "What has been said . . . does not prove that the superior ability of the early developers stems from their greater physical maturity. It is more likely that the explanation lies in the high proportion of these girls who come from small families and who are only children. It is most suggestive that at 11 years the measured ability of only children bears no relationship to their stage of sexual development."

The relationship between measured ability and family size has been too well demonstrated to require detailed consideration here. Nisbet (1953) summarised studies which showed a negative correlation between children's intelligence test scores and the size of their families. He showed that " . . . Part (but not all) of this negative correlation may be attributed to the environmental influence of the size of the family on verbal ability and through it, on general mental development." It will, therefore, be more pronounced when tests are used which tap verbal ability, a hypothesis confirmed by Nisbet and Entwistle (1967). It is an influence which appears to exercise its maximum effect before the age of 7 and probably before the age of 2 (Honzik, 1967).

The relationship between rate of growth and family size has been noted and discussed by Tanner (1962) and Douglas (1964). It holds over all socio-economic groups and attempts have been made to account for it in terms of differences in nutritional standards between large and small families. Douglas points out that such an explanation cannot account for all the known facts however, and the determining factors remain hypothetical. Whatever they may be, it is probable that they affect also the general standard of child care and through it, cognitive development ; at any rate there is a *prima facie* case for considering the interaction of the three variables, rate of growth, family size and educational progress as they appear at adolescence.

II.—PROCEDURES.

The present investigation is based on a sample of 1,064 girls in three West Riding comprehensive schools between the ages of 11½ and 15½. In 1964 they completed a carefully designed questionnaire which gave information about the age of menarche. For the older girls this gave a practically complete record and the remainder were followed up until the records were complete for all girls in the first four years of secondary schooling. The techniques used and the results are described in greater detail in a previous paper (Poppleton and Brown, 1966). The sample covers the whole range of ability and the schools are situated fairly close together in a mixed rural-mining community of comparatively low socio-economic status. All the schools are coeducational but only the girls were chosen as subjects because of the difficulty of obtaining accurate information about age of puberty in the case of boys.

In March, 1964, when records of menarcheal status were begun, the girls also completed Test AH4, Parts 1 and 2 of which give measures of verbal and non-verbal intelligence. The yearly examination results in English and Mathematics were made available and were scaled on Part 1 of the test as this includes both verbal and numerical items. Age allowances were calculated for each year group separately and the results converted to normalised T-scores. Information was also obtained concerning family size and father's occupation which was the only index of social class which was readily available in this case. As the sample was faily homogeneous in this respect, a threefold classification was adopted of professional and managerial, non-manual and manual occupations. The girls were classified into four maturity groups of equal size on the basis of age at menarche, each year group being divided at the quartiles and the median.

Statistical Analysis.

A complex analysis of variance based on all variables was not feasible as there would not have been equal numbers in all cells. Consequently, differences between the maturity groups on the measures of intelligence and attainment were examined by a one-way analysis of variance. A previous investigation (Poppleton, 1966) had shown that these groups did not differ with respect to social class; and so, where significant differences were obtained, an analysis of covariance was carried out controlling family size.

Correlation coefficients were next calculated between all the variables in the study; age at puberty, father's occupation, family size, school, AH4 Parts 1 and 2 and English and Mathematics attainment. The resulting matrices were submitted to a principal components factor analysis with varimax orthogonal rotations.

III.—RESULTS.

The mean family size and age at menarche of the sample year groups are shown in Table 1.

TABLE 1

MEAN FAMILY SIZE AND AGE AT MENARCHE OF SAMPLE YEAR GROUPS.

		Family Size		Age at Menarche	
	N	M	S.D.	M	S.D.
1st Year (11 : 6—12 : 5)	280	3·25	1·96	13·05	0·91
2nd Year (12 : 6—13 : 5)	280	3·24	1·74	13·07	0·90
3rd Year (13 : 6—14 : 5)	252	3·14	1·83	13·28	0·91
4th Year (14 : 6—15 : 5)	252	3·20	1·76	13·31	0·95

TABLE 2

MEAN T-SCORES OF FOUR MATURITY GROUPS.

	Early	Low Av.	High Av.	Late		
Age	Before 12 : 6	12 : 7—13 : 4	13 : 5—14 : 2	Over 14 : 2	F-ratio	Adjusted F-ratio
1st Year:						
AH4 Part 1	51·9	49·8	49·7	49·9	1·26	
Part 2	52·7	49·3	49·1	49·0	1·82	
English	51·4	49·8	50·2	49·9	1·01	
Mathematics	51·9	49·0	49·9	49·2	1·41	
2nd Year:						
AH4 Part 1	51·5	50·3	50·4	47·8	1·78	
Part 2	51·2	50·4	51·1	48·2	1·43	
English	52·4	50·7	49·4	47·3	2·98*	1·04
Mathematics	51·2	51·1	48·5	48·4	1·32	
3rd Year:						
AH4 Part 1	51·4	51·4	50·7	49·7	1·25	
Part 2	50·7	51·0	50·5	48·6	1·82	
English	50·3	51·0	49·2	49·4	1·71	
Mathematics	51·5	50·7	49·8	49·5	1·55	
4th Year:						
AH4 Part 1	51·9	50·2	50·3	48·0	1·69	
Part 2	50·5	50·0	51·4	48·2	1·12	
English	51·0	51·1	50·0	47·8	1·51	
Mathematics	50·8	50·7	49·7	48·1	1·00	

* F for values of d.f. $>$ 120 and 3 at $p < \cdot 05 = 2\cdot 60$.

Table 2 shows the mean scores of the four maturity groups in each year on each of the four educational measures. Looking along the rows, there is a consistent tendency for early maturers to be ahead of later maturers in all cases, though in only one case is the difference sufficiently large to achieve a significant F-ratio. This occurs in the English results of the second year girls and accordingly, an analysis of covariance controlling family size was performed on these scores and the adjusted F-ratio was well below the level required for significance. In spite of the higher F-ratio obtained on verbal measures in the second year, the trends in all years are not particularly associated with superiority on verbal measures only.

Looking down the columns, there are only small shifts in the mean scores from year to year. The early maturers maintain much the same level of performances throughout the years, the average groups show slight improvement between the first and second years and maintain this level, while the late maturers tended to lose a little ground and score lower in the fourth year. There is no evidence that they ' spurt ' during these years and catch up on their earlier maturing age-mates.

There is little doubt that if all the means in Table 2 were adjusted for differences in family size, the differences between the scores of the extreme maturity groups would be considerably narrowed and may well disappear altogether. Table 3 shows a breakdown of the family size of each maturity group. There is a clear tendency for early maturing girls to come from smaller families than late maturing girls, though an analysis of variance shows the differences to be significant in the first and second years only. It might be argued that family size cannot be regarded as complete until these girls have reached the age of 14 or 15, but according to Nisbet (1953), family size can be regarded as reasonably complete by the time the eldest child is 11 or 12. It is unlikely that we are seeing here the effect of incomplete family size as the means of all four samples are reasonably uniform and, in fact, the total mean is higher in the first year than the fourth (3.24 as compared with 3·20).

TABLE 3

MEAN FAMILY SIZE OF MATURITY GROUPS.

	Early	Low Av.	High Av.	Late	F-ratio
1st Year	2·80	2·99	3·63	3·59	3·26*
2nd Year	2·76	3·23	3·26	3·71	3·59*
3rd Year	2·80	2·90	3·30	3·60	2·39
4th Year	3·02	3·14	3·37	3·30	2·10

* F-ratio significant at $p < \cdot 05$.

Further evidence regarding the interaction of all the variables included in this study is offered by the pattern of correlations obtained. Correlations between age at puberty and the educational variables were all small and non-significant but those between age at puberty and family size were significant ($p < \cdot 05$) for the first two years. Father's occupation showed no significant relationship with age at puberty or family size but showed significant correlations ($p < \cdot 01$) with all the educational variables. Family size correlated negatively and significantly with all the educational variables and in the case of both verbal and non-verbal intelligence showed a tendency to increase up to the age of 15.

The results of the factor analysis are shown in Table 4.

TABLE 4

FACTOR LOADINGS: VARIMAX ORTHOGONAL ROTATION.

	N=280 1st Year		N=280 2nd Year		M=252 3rd Year		N=210 4th Year	
	I	II	I	II	I	II	I	II
1. Age at Puberty	007	–808	064	–852	101	–870	–016	–778
2. Father's Occupation	–442	–229	–292	–297	–430	–354	–134	–716
3. Family Size	–208	–610	–233	–598	–258	–496	–425	–040
4. School	–667	–046	–673	–020	–698	–112	–777	027
5. Intelligence Verbal	803	–169	906	201	916	135	899	144
6. Intelligence Non-Verbal	807	–158	821	162	821	146	785	107
7. English attainment	876	–195	882	248	899	177	841	161
8. Maths. attainment	895	–187	909	144	914	090	891	071
% Variance	46·49	15·80	46·18	16·50	48·79	15·03	46·61	14·79
	61·57		62·68		63·82		61·40	

(Decimal points omitted)

The analysis was undertaken for each year group separately and yielded in each case the two orthogonal factors which together accounted for over 60 per cent of the variance. Factor I accounted for 46–49 per cent of the total variance and Factor II for 14–17 per cent, the contribution varying slightly between years. Table 4 shows the factor loadings of the eight variables. (The sample size of the fourth year was reduced because information regarding father's occupation could not be supplied by one school.)

Factor I.—The major part of the variance on this factor is contributed by the measures of intelligence and attainment which together account for between 78 and 82 per cent of the total variance. Age at puberty has near zero loadings throughout and the remaining three environmental variables make up the remaining contribution. This seems therefore to be an educational-social factor.

TABLE 5A

PERCENTAGE CONTRIBUTION TO THE SOCIAL VARIANCE ON FACTOR I (VARIABLES 2-4).

Year	1	2	3	4
School	65·1	76·5	65·9	75·3
Father's Occupation	28·6	14·4	25·0	2·2
Family Size	6·3	9·2	9·0	22·5

The contribution of the environmental variables to the total social variance on this factor is shown in Table 5A. School contributes the major part of the variance while, for the first three years, occupational category exercises a greater influence than family size. The correlations on which these figures are based are small and subject to sampling fluctuations which would account for variations in the size of the contributions. The reversal in rank order between father's occupation and family size in the fourth year is puzzling. It has already been mentioned that information about occupational category was not available at one school and, in general, greater difficulty was experienced in collecting accurate data on the fourth year sample than in other cases. There is some reason to suspect its reliability and only a further investigation of this age group would clarify the issue.

The loadings on the verbal part of the test are almost all higher than on the non-verbal part. Mathematics has the highest loadings in the first and second years and the next highest in the third and fourth ; it has a higher loading than English throughout. The differences are not great, but the pattern is reasonably consistent from year to year.

Factor II.—Over half the variance on this factor is contributed by age at puberty which suggests that it can be identified as a developmental factor (Table 5B). The contribution of family size in the first three years confirms the relationship of these variables. Once again, the reversal of this with the socio-economic variable in the fourth year is noticeable and must be attributed to the explanation given above. The contribution from father's occupation is small but increases in importance. The loadings of the educational variables on this factor are all small, but generally are higher on English attainment and AH4 Part 1. This factor therefore suggests a link between verbal achievement, rate of maturation and family size with economic factors playing a small but increasingly important role.

TABLE 5B

PERCENTAGE CONTRIBUTION TO THE TOTAL VARIANCE ON FACTOR II (VARIABLES 1-3).

Year	1	2	3	4
Age at puberty	54·1	54·9	62·9	51·2
Father's Occupation	4·3	6·7	10·4	43·4
Family Size	30·9	27·1	20·5	0·2
	89·3	88·7	93·8	94·8

IV.—DISCUSSION.

Studies of the factors entering into ability and educational progress have tended to concentrate wholly on either environmental or developmental influences and, as the introduction showed, certain facts relating to two of these influences (age of puberty and family size) have been well established when they are considered independently. The main aim of this study has been to explore the mode and extent of the interaction of these factors as they appear in adolescence.

The results showed that differences in ability and attainment between groups of girls at differing stages of physical maturity were negligible when family size was taken into account. Supportive evidence came from the factor analysis in that age at puberty made no significant contribution to Factor 1 which took up most of the educational variance. It was, however, associated with the small remaining part of the educational variance on Factor 2 in conjunction with family size.

The influence of family size is thus expressed in two ways : firstly, through its contribution to the social variance on Factor 1. Girls from large families in this sample tended to be in non-selective schools and to have parents in the lower occupational categories. They also tended to a lower level of performance in mathematics and do less well on a test of verbal intelligence. Secondly, however, it is strongly associated with age at puberty in Factor 2, showing its influence on the rate of both physical and mental development which is reflected in poorer performance in English as well as verbal ability. Poor verbal ability seems therefore to be associated with both environmental and developmental factors. Non-verbal ability is also affected though to a slightly less extent.

In conclusion, it would appear that family size is a more important variable than age at puberty in producing significant differences in ability and attainment. The influence of both is consistent over the years so that there is no case for introducing developmental weightings at any point during the early years of secondary schooling. The association between family size and the socio-economic variable however, comprises a variable influence which can be augmented or diminished by the type of schooling received. It is by far the most important and susceptible to remedial measures.

ACKNOWLEDGEMENTS.—The author wishes to acknowledge the help received from Professor F. W. Warburton and the Manchester School of Education computing staff for help with the factor analysis. The research was made possible by a grant from the University of Sheffield Research Fund Committee.

V.—REFERENCES.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London : MacGibbon and Kee.

HONZIK, M. P. (1967). Environmental correlates of mental growth : prediction from the family setting at 21 months. *Child Developm.*, 38, 337-365.

NISBET, J. D. (1953). *Family Environment.* Eugenics Society Occasional Paper No. 8. London : Cassell.

NISBET, J. D., and ENTWISTLE, N. J. (1966). *The Age of Transfer to Secondary Education.* Scottish Council for Research in Education. Univ. of London Press.

NISBET, J. D., and ENTWISTLE, N. J. (1967). Intelligence and family size, 1949-1965. *Brit. J. Educ. Psychol.*, 37, 188-193.

PLOWDEN, B. (1966). *Children and their Primary Schools.* Report of Central Advisory Council for Education. London : H.M.S.O.

POPPLETON, P. K. (1966). *The Inntelligece abd Attainments of Premenarcheal and Postmenarcheal Girls.* Unpublished paper read to B.P.S. (Education Section) Conference.

POPPLETON. P. K. and BROWN, P. E. (1966). The secular trend in puberty : has stability been achieved ? *Brit. J. Educ. Psychol.*, 36, 97-101.

TANNER, J. M. (1961). *Education and Physical Growth.* Univ. of London Press.

TANNER, J. M. (1962). *Growth at Adolescence* (2nd Edit.). Oxford : Blackwell.

(*Manuscript received* 26*th February*, 1968)

ANXIETY AND EDUCATIONAL ACHIEVEMENT

By BARRY P. FROST

(*Department of Educational Psychology, University of Calgary, Alberta*)

SUMMARY. This paper reports a study of the relations between anxiety and educational achievement in 310 11-year-old junior school pupils (170 boys, 140 girls). The instruments used were the Frost Self Description Questionnaire (which includes seven anxiety scales and a Denial scale), the Junior Eysenck Personality Inventory, Progressive Matrices, Crichton Vocabulary and three NFER tests of reading and arithmetic. Product-moment, additive multiple regression and Principal Components (with Varimax rotation) analyses are reported. The results demonstrate (*a*) a complex inter-relationship of types of anxiety, types of achievement and sex of subject; (*b*) that no separate 'general' and 'school' anxiety factors emerge; (*c*) the necessity of using a 'denial' scale, and (*d*) absence of correlation between Eysenck's Neuroticism Scale and school achievement. A theoretical discussion of 'General Emotionality' and anxiety led to the conclusion that anxiety questionnaires are likely to measure (1) the likelihood of being threatened by the external world, and (2) a specific way of reacting to such threat.

I.—INTRODUCTION.

CURRENT theories dealing with the relations of anxiety to learning or educational achievement, e.g., those developed by Eysenck (1955), Spence (1958), and Sarason, *et al.* (1960), leave much to be desired because of their lack of specificity. There is, in these theoretical treatments, a lack of specificity both with respect to type of anxiety and educational achievement criteria. More empirical studies such as those of Phillips (1962) have attempted to be more sophisticated with respect to the differentiation of educational variables (and with respect to modulator variables such as age, sex and social class) but they lack a distinctive theoretical framework.

The anxiety scales currently in use, such as Taylor's Manifest Anxiety Scale (MAS), the Children's Manifest Anxiety Scale (CMAS), Sarason's General Anxiety Scale for Children (GASC) and Test Anxiety Scale for Children (TASC), are collections of items, the denotation and connotation of which are variable. Few attempts have been made to analyse the items in terms of differential measurement as Dixon (1955) has done for the anxiety items from the Tavistock Self-Assessment Inventory. Dixon intercorrelated the twenty-six 'phobic' items from this inventory and factor-analysed the resultant matrix. He found two factors which he named 'separation' anxiety and 'castration' anxiety. Similarly he analysed the 26 'social anxiety' items from the inventory. This time he obtained a 'general social anxiety' factor and four group factors—'social timidity,' 'fear of loss of control,' 'fear of exhibitionism,' and 'fear of revealing inferiority.'

O'Connor, *et al.* (1956) factor-analysed the MAS and found five factors—(1) chronic anxiety or worry; (2) increased physiological activity; (3) sleep disturbances associated with inner strain; (4) sense of personal inadequacy; and (5) motor tension.

Following up this study, Bendig (1958) found the following relations between the E and B Scales of the Maudsley Personality Inventory (MPI) and the five factor scales of the MAS as derived by O'Connor, *et al.*

Extraversion had essentially zero correlations with all but one of the factor scales. The exception was 'Sense of Personal Inadequacy,' and the correlation here was −0·51. Neuroticism had small positive correlations with all scales except one. 'Increased Physiological Reactivity' correlated −0·49 with Neuroticism.

Dunn (1964) gave a slightly modified version of TASC to 633 pupils aged 9 to 12, in five schools in an upper-middle-class urban community, and factor-analysed the results. He obtained four factors which he named: (1) Test Anxiety; (2) Generalized School Anxiety; (3) Recitation Anxiety; and (4) Physiological Arousal in Anticipated Recitation Situation.

No attempt has been made to delineate sub-scales of any of the children's anxiety tests. As a first step in this direction the writer placed items from the CMAS, GASC and TASC (on face validity) into the separate scales: (1) Autonomic Anxiety; (2) Concentration Anxiety; (3) Worries and Fears; (4) Social Anxiety; (5) Separation Anxiety; (6) Body Damage Anxiety; and (7) Non-specific Anxiety, so termed because of their item content. The items were, when necessary, re-written to suit the age group involved. All of the items in the first four scales refer to the school situation and collectively form the 'School Anxiety' scale. The last three scales together form the 'General Anxiety' scale. Items, largely derived from the Tavistock Self-Assessment Inventory, were re-written and added so that each of the scales has ten items (an exception being the Non-Specific Anxiety scale which has twelve items).

These anxiety scales together with aggression, affiliation, assertiveness, submission and denial scales plus buffer items form the Frost Self Description Questionnaire (Frost, 1967). The Denial Scale was designed to test the tendency of the subject to deny that behavior commonly found in a normal population, or emotions commonly experienced, are in fact committed or experienced by him. It is assumed that such a measure will bear some relation to the defence mechanism of 'Denial.'

II.—PLAN OF THE INVESTIGATION.

The Frost Self Description Questionnaire was given to 170 boys and 140 girls, aged between 10 years 11 months and 11 years 9 months, members of the fourth classes of ten different junior schools in London. The girls had a mean age of 11.32 years, S.D. 3·71 months; the boys, a mean age of 11·29 years, S.D. 3·27 months. In seven out of the ten schools, all of the fourth class children were tested; in the eighth, a 50 per cent sample was tested, the sample being selected by taking every second name on the register; in the ninth, a one-third sample was taken in the same fashion; in the tenth, only one class was tested as the other was composed largely of migrant children.

The papers of migrant children in the classes tested were discarded and not used in any way in this study. Also eliminated were children who were absent for all or part of the testing. These children (27 boys and 27 girls), according to the descriptions given by their head teachers, were not in the main noticeably different from their peers. Six of them were behavior problems as far as the school was concerned. After testing was completed a further 15 children were discarded as they were either older or younger than the age-range selected. In total the wastage from the selected fourth class native-born children was approximately 18 per cent (or 15 per cent of the age-range chosen). There does not seem to be any reason to consider the subjects of the investigation as other than a random sample of fourth class children (with the exception that there were no upper class nor educationally sub-normal children contained in the sample).

The schools were junior schools in a North London borough which is characterized by a range of ecological conditions, from wide parkland to terraced streets with factories and/or shops. The schools range in social class background from largely middle class to virtually completely working class.

The testing involved in the present study was that carried out for a larger study (Frost, 1967) and involved apart from the Frost Self Description Questionnaire, the Junior Eysenck Personality Inventory, Raven's Progressive Matrices, Crichton Vocabulary Scale, NFER Reading Comprehension Test N.S.6, NFER Mechanical Arithmetic Test N.S.10 and NFER Problem Arithmetic Test N.S. 53.

The testing was carried out either in the children's own classroom or in a vacant classroom by the writer, no teacher being present.

III.—STATEMENT OF RESULTS AND DISCUSSION.

In the present report reference will be made only to the Anxiety scales, the Denial Scale, the Neuroticism Scale of the Junior Eysenck Personality Inventory, Raven's Matrices, Crichton Vocabulary Scale and the three educational achievement criteria.

(1) *Correlations.*

Product-moment correlation coefficients for all variables were calculated, separate matrices being derived for boys and girls respectively. These matrices are shown in Table 1. The results of these analyses show that there is a complex inter-relationship of types of anxiety, types of achievement and sex of the subject.

To take sex of subject first, it can be seen that with respect to Concentration Anxiety, the product-moment correlations are significant though small in the case of the boys, but not significant in the case of the girls (reading excepted). Presumably concentration anxiety is more potent in boys than girls.

Contrariwise, the correlations between Separation Anxiety and the three educational criteria are all essentially zero in the boys' matrix but significant in the girls'. There are also some differences between the sexes in the correlations between the three criteria and Body Damage Anxiety but the size of the differences is not sufficient for us to believe in true differences with any degree of certainty.

Contrary to expectation there are few major differences in the relationships of the different anxiety scales to the three achievement criteria. The outstanding exception to this situation is Social Anxiety. With the exception of Reading in girls, Social Anxiety is not significantly related to any achievement variable in either sex.

Sarason (1957) and others, e.g., Alpert (1958), Cox (1960), have claimed that 'test' or 'school' anxiety is more specifically related to educational achievement than 'general' anxiety. The present results do not support this contention.

The coefficients in the boys' matrix do not show a consistent significant difference between the 'school' and 'general' anxiety scales *vis-a-vis* their relations with the three educational criteria. In the girls' matrix the correlations between the 'general' scale and the three criteria are higher than those between the 'school' scale and the criteria.

TABLE 1

INTER-CORRELATION MATRIX
(Decimals Omitted)
Girls (N=140) *above diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Autonomic Anxiety	.	56	66	63	60	47	70	−11	85	72	84	59	−09	−27	−31	00	−14
2. Concentration Anxiety	59	.	52	54	37	31	64	−18	77	55	71	59	−10	−15	−16	−01	−10
3. Worries and Fears	69	52	.	63	52	50	68	−15	85	68	82	59	−04	−24	−26	−04	−16
4. Social Anxiety	56	38	55	.	57	44	73	−21	84	71	82	59	−00	−15	−16	02	−06
5. Separation Anxiety	46	37	48	46	.	53	61	01	61	85	75	53	−14	−33	−36	−22	−29
6. Body Damage Anxiety	40	39	38	38	35	.	50	−02	53	77	66	49	−11	−28	−32	−17	−24
7. Non-Specific Anxiety	62	55	64	54	50	40	.	−09	83	87	89	79	−09	−29	−32	−12	−20
8. Denial	−06	−13	−17	−17	−20	−01	−12	.	−20	−04	−14	−22	−30	−47	−46	−44	−41
9. School Anxiety (Variables 1-4)	87	76	85	76	54	48	73	−16	.	80	96	71	−06	−24	−27	−00	−13
10. General Anxiety (Variables 5-7)	64	56	66	59	78	70	83	−14	76	.	93	74	−14	−37	−40	−20	−29
11. Total Anxiety (Variables 1-7)	83	72	82	74	68	60	82	−17	96	91	.	77	−10	−30	−34	−09	−21
12. Neuroticism (Jn. E.P.I.)	50	55	53	45	49	38	66	−27	64	66	69	.	−06	−20	−20	−03	−11
13. Raven's Matrices	−13	−08	−03	−02	04	−02	−06	−23	−08	−01	−06	02	.	60	63	56	67
14. Crichton Vocabulary	−19	−22	−11	−08	−07	−16	−17	−27	−18	−16	−18	−02	58	.	88	61	73
15. Reading Comprehension	−16	−23	−09	−05	−06	−17	−21	−26	−16	−18	−18	−09	53	84	.	66	77
16. Mechanical Arithmetic	−20	−18	−09	−13	04	−18	−13	−23	−19	−11	−17	−05	55	61	65	.	81
17. Problem Arithmetic	−20	−24	−11	−08	−03	−16	−17	−19	−19	−14	−19	−08	65	75	75	82	.

Boys (N=170)** below diagonal

* An r of 0·14 is significant at the 5 per cent level (one tail test).

** An r of 0·13 is significant at the 5 per cent level (one tail test).

Another point of interest is the essentially zero relationship between the Eysenck Neuroticism Scale and the three educational criteria (reading in girls being the sole exception) as compared with the significant coefficients between several of the specific anxiety scales and the criteria.

(2) *Factor Analysis.*

In the larger study referred to previously (Frost, 1967) the product-moment inter-correlation matrices for all independent variables were subjected to a Principal Components Analysis with Varimax Rotation. From these Varimax analyses the loadings of the variables dealt with in the present paper have been extracted and are present in Table 2.

TABLE 2

ROTATED (VARIMAX) FACTORS.
(Loadings 0·30 and above)

	Boys				Girls		
Scale/Factor	I	II	III	IV	I	II	III
1. Autonomic Anxiety	69			32	81		
2. Concentration Anxiety	52			44	68		
3. Worries and Fears	72				80		
4. Social Anxiety	75				82		
5. Separation Anxiety	69				74		
6. Body Damage Anxiety	39			48	63		
7. Non-Specific Anxiety	82				85		
8. Denial		−30	62			−51	53
9. Neuroticism	72				80		
10. Progressive Matrices		77				75	
11. Vocabulary		85				81	
12. Reading		85				85	
13. Mechanical Arithmetic		83				86	
14. Problem Arithmetic		90				89	

The First Factor: 'General Emotionality.'

The first factor in both the boys' and the girls' matrices is clearly akin to the factor found by Burt (1915, 1965) and termed, by him, 'general emotionality.' It *positively* loads *all* of the anxiety scales and Neuroticism. It also, in the larger study (see Frost, 1967), loads the intropunitive scales, Self Criticism and Guilt, as well as Submission.

This factor of 'General Emotionality' (Burt, 1915) or 'General Reactivity' (Adcock, 1965) clearly is analogous to Eysenck's 'Neuroticism' factor (Eysenck, 1957) and Cattell's second-order factor termed 'Anxiety' (Cattell, *et al.*, 1962). However, the relatively neutral term suggested by Burt is to be preferred to either 'Neuroticism' or 'Anxiety.' The former term is correctly criticised by Burt as implying a neurotic state or instability which is not necessarily the case, "the degree of a person's instability does not always correspond precisely with his degree of general emotionality. . . a better estimate of a person's instability may be obtained by taking the ratio of his general emotionality to his general intelligence" (Burt, 1965).

Adcock (1965) argues that both 'Neuroticism' and Cattell's 'Anxiety' are equivalent to 'General Reactivity' and (like Burt) that more is implied by the former terms than the latter. The factor is also presumably analogous to 'Drive' in the Spence-Taylor theory, i.e., a generalized level of reactivity. If it is related to a general level of reactivity it is also, presumably, related to 'Arousal' (Duffy, 1957) or 'Activation' (Hebb, 1955; Malmo, 1959). However, Welford (1965) links 'neuroticism' not to arousal level as such (which he sees as related to introversion-extraversion) but to "change in arousal level with stress," and cites evidence from Davies, *et al.* (1963). The evidence cited is based, however, on inadequate experimental data. The investigation was carried out on 35 subjects only (a mixed group of neurotics and normals), and few of the correlation coefficients were significant. Furthermore, the coefficient largest in magnitude was that between the N and E scales of the MPI (−0·50), a somewhat surprising datum in an experiment which is supposed to link E to level of arousal and N to change in level. As these dimensions (E and N) are alleged to be orthogonal, the result leads one to suspect that the sample was biased in some way.

General Emotionality, Intropunitive Aggression and Submission.

The Frost Self Description Questionnaire, in addition to the anxiety scales, includes five aggression scales, an Assertiveness Scale and a Submission Scale. The measures of intropunitive aggression (Self-Criticism and Guilt) and Submission are loaded by the general emotionality factor whereas the measures of extrapunitive aggression (Overt Aggression, Criticism of Others, Projection) and the Assertiveness Scale are not.

If we take the view that general emotionality is a reflection of resting arousal level, then presumably, the higher the level of arousal the greater the degree of intropunitive aggression. It seems strange that a high level of arousal would not also result in a high degree of extrapunitive aggression.

If, on the other hand, we take Welford's view that general emotionality is a function of change of reactivity (or arousal) with stress, then the greater the change in arousal with stress the greater the degree of intropunitive aggression. Extrapunitive aggression, presumably, would have no relationship to change in arousal level under conditions of stress, which seems contrary to the evidence about aggressive responses to frustration.

Another view which might be taken of general emotionality is that it is a form of psychological reaction to a given level of activation, a form of reaction characterized by a tendency to deprecate the self and a failure to take action. Cattell (1963) says that "when a person grapples with a difficulty, he shows the stress response; when he retreats and attempts escapist solutions, he shows anxiety."

Generally speaking, however, it makes more sense to treat general emotionality as Burt treats it, i.e., as analogous to McDougall's "common fund of emotional energy," i.e., autonomic level. We know that depressives are highly aroused (Stern and McDonald, 1965) and this fits with our evidence that the intropunitive scales and Submission are positively loaded by the general emotionality factor. Similarly, eight separate studies found depressives to have the highest scores on Neuroticism (Ingham and Robinson, 1964) and Hope (1963) noted that "the more neurotic a subject is the more aggression he has and the more he turns it against himself." However, it would seem that there are at least six possible interpretations of the factor of 'general emotionality': (1) degree of acquisition of anxiety linked signs; (2) degree of autonomic response to threat; (3) change in reactivity under conditions of stress; (4) likelihood of anxiety response to threat; (5) psychological reaction to a high arousal level of a self-deprecatory and action-avoiding type; (6) likelihood of being threatened by the external world. Of these alternatives, the writer feels that the last is the only one which indisputably follows from the type of data used in this study.

General Emotionality and Achievement

It seems that 'general emotionality' is orthogonal to 'intellectual achievement' whatever the relations of the different forms of anxiety may be to such achievement. It is appropriate here to recall Burt's comment, that instability is related to the ratio of general emotionality to general intelligence.

Anxiety: Unitary or Plural?

While the greater part of the variance of each anxiety scale was accounted for by the 'general emotionality' factor, there seems to be a small degree of specificity of function, some aspects of which will be dealt with in the multiple regression analysis section. There was no evidence in our factor analysis to support the notion of separate 'school' (or 'test') and 'general' anxiety factors.

Though there is no evidence for the existence of these particular factors it seems necessary to distinguish between 'general emotionality' and 'anxiety.' Anxiety, as a concept, implies *direction* as well as *level* whereas 'general emotionality' may imply level of arousal without any specific directionality (other than, possibly, a generalized tendency to be concerned or experience threat). If we accept the 'anticipatory' or 'signalling' concept of anxiety, then, presumably, signalling may occur in one setting, e.g., school, without it occurring in another, e.g., home. On the other hand, if degree of general emotionality implies level of arousal, which in turn implies conditionability or ease of acquisition of negative signs, then it would seem that the greater the degree of general emotionality, the greater the likelihood of 'signal anxiety' irrespective of setting. None of the six possible interpretations of general emotionality referred to above runs counter to the idea of anxiety as specific signal, but, if we accept the sixth version then each anxiety scale can be said to contain two components, (1) the likelihood of being threatened by the external world, (2) a specific way of reacting to such threat. Such an analysis does not contradict the basic idea that anxiety is an experience dependent upon a given level of arousal; it only calls into question the notion that the 'general emotionality' factor measures level of arousal.

Multiple Regression Analysis.

Six additive multiple regression analyses were made of the data. Boys and girls were taken separately for each of the three criteria—Reading, Mechanical Arithmetic and Problem Arithmetic. The program used allowed of the testing of the effect on the multiple R of adding each predictor variable. While the Principal Components analysis was performed to throw light on the nature and relative independence of the dimensions concerned in the investigation, the Multiple Regression analysis was carried out to show to what extent the variance of the criteria measures—Reading, Mechanical Arithmetic and Problem Arithmetic could be accounted for by the anxiety variables. Table 3 shows the results of this analysis.

This table shows that some anxiety scales contribute to an increase in R and some do not. The specific scales which do contribute vary according to the criteria and the sex of the subjects.

It is clear in all analyses, however, that Denial contributes considerably to the prediction of the criterion variance. This finding supports the thesis proposed by previous investigators—that a denial or 'lie' scale be included in the battery. The writer differs from these previous investigators in believing that such scales measure a generalized and probably unconscious refusal to face reality rather than conscious lying. In another paper (Frost, 1968) this problem of denial will be dealt with in more detail.

TABLE 3

MULTILPLE REGRESSION ANALYSIS.

Criteria :	Reading		Mechanical Arithmetic		Problem Arithmetic	
	Girls	Boys	Girls	Boys	Girls	Boys
Predictors	R*	R	R	R	R	R
1. Autonomic Anxiety	(–·29)	(–·16)	(·00)	(–·20)	(–·14)	(–·20)
2. Concentration Anxiety	·29	·23	·01	·22	·13	·26
3. Worries and Fears	·30	·24	·08	·24	·17	·28
4. Social Anxiety	·32	·24	·10	·24	·18	·28
5. Separation Anxiety	·40	·24	·32	·29	·36	·29
6. Body-Damage Anxiety	·43	·27	·34	·31	·38	·29
7. Non-Specific Anxiety	·44	·30	·36	·31	·39	·30
8. Denial	·64	·41	·61	·38	·54	·36

* NOTE.—Each R is that obtained by *adding* the specific variable to the Multiple Regression formula. The figure for Autonomic Anxiety is the Pearson r.

The multiple regression analyses clearly show that, using personality variables alone without any contribution from intelligence tests, a considerable proportion of the criterion variance can be predicted. In the present investigation R varied between 0·36 and 0·64, i.e., between 13 and 41 per cent of the criterion variance was predicted by the anxiety and denial scales.

A partial replication of this study using the same personality questionnaire with Canadian children has been carried out and the data are being analysed at present. Following on a factor analysis of the items of the questionnaire a revised version has been prepared and is being used in a further study of Canadian children. This revised version of the questionnaire is being translated into Spanish and Japanese for use in Mexico, parts of South America and Japan. The writer is planning collaboration with colleagues in these countries, with respect to cross-cultural studies of anxiety and aggression and their relation to scholastic achievement.

IV.—CONCLUSIONS.

It is clear that, although there is a great deal of commonality in anxiety scales, which the writer, following Burt, has termed 'general emotionality,' there is also much specificity. This specificity is important with respect to educational achievement as, it seems, some forms of anxiety do not relate to such achievement. It is also clear that there is a complex inter-relationship of type of anxiety, type of achievement and sex of subject. It appears that there is a considerable correlation between 'denial' and educational achievement and that this aspect needs further research. While there are several possibilities with respect to the nature of 'general emotionality' and anxiety as measured by a questionnaire, it is likely that they refer to: (1) the likelihood of being threatened by the external world, and (2) a specific way of reacting to such threat.

Acknowledgements.—The research reported in this paper was aided by a grant from the Irwin Fund of the University of London. All Product-Moment correlations, Principal Components and Varimax analyses were calculated on the IBM 1620 computer at the University of Calgary. I am indebted to Professor David Flathman for the program used and the supervision of the computation. The Multiple Regression analyses were calculated on the IBM 360 computer at the University of Calgary. I would like to thank Dr. Herbert Hallworth for the program used and Mr. K. E. Nichols for the supervision of the computation.

IV.—References.

Adcock, C. J. (1965). A comparison of the concepts of Cattell and Eysenck. *Brit. J. Educ. Psychol.*, 35, 90-97.

Alpert, R. (1958). Anxiety in academic achievement situations : its measurement and relation to aptitude. *Dissert. Abstr.*, 18, 643.

Bendig, A. W. (1958). Identification of item factor patterns within the Manifest Anxiety Scale. *J. Consult. Psychol.*, 22, 158.

Burt, C. (1915). The general and specific factors underlying the primary emotions. *Brit. Ass. Ann. Rep.*, 84, 694-696.

Burt, C. (1965). Factorial studies of personality and their bearing on the work of the teacher. *Brit. J. Educ. Psychol.*, 35, 368-378.

Cattell, R. B., Scheier, J. H., and Lorr, M. (1962). Recent advances in the measurement of anxiety, neuroticism and the psychotic syndromes. *Ann. N.Y. Acad. Sci.*, 93, 813-856.

Cattell, R. B. (1963). The nature and measurement of anxiety. *Scientific American*, Reprint No. 475.

Cox, F. N. (1960). Correlates of general and test anxiety in children. *Aust. J. Psychol.*, 12, 169-177.

Davies, M. N., Claridge, G. S. and Wawman, R. J. (1963). Sedation threshold, autonomic lability and the excitation : inhibition theory of personality : III—The blood pressure response to an adrenaline antagonist as a measure of autonomic lability. *Brit. J. Psychiat.*, 109, 558-567.

Dixon, J. J. (1955). *An Investigation of Some Neurotic Anxieties.* Unpub. Ph.D. thesis, Univ. of London.

Duffy, Elizabeth (1957). The psychological significance of the concept of ' Arousal ' or ' Activation.' *Psychol. Rev.*, 64, 265-275.

Dunn, J. A. (1964). Factor structure of the Test Anxiety Scale for Children. *J. Consult. Psychol.*, 28, 92.

Eysenck, H. J. (1955). A dynamic theory of anxiety and hysteria. *J. Ment. Sci.*, 101, 28-51.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London : Routledge and Kegan Paul, 1957.

Frost, B. P. (1967). *Some Personality Conditions of Educational Achievement in Children in the Fourth Class of Junior School.* Unpubl. Ph.D. thesis, Univ. of London.

Frost, B. P. (1968). Denial and scholastic achievement (in preparation).

Hebb, D. O. (1955). Drive and the C.N.S. (Conceptual Nervous System). *Psychol. Rev.*, 62, 243-254.

Hope, K. (1963). *The Structure of Hostility Among Normal and Neurotic Persons.* Unpubl. Ph.D. thesis, Univ. of London.

Ingham, J. G., and Robinson, J. O. (1964). Personality in the diagnosis of hysteria. *Brit. J. Psychol.*, 55, 276-284.

Malmo, R. (1959). Activation : a neuropsychological dimension. *Psychol. Rev.*, 66, 367, 386.

O'Connor, J. P., Lorr, M., and Stafford, J. W. (1956). Some patterns of manifest anxiety. *J. Clin. Psychol.*, 12, 160-163.

Phillips, B. N. (1962). Sex, social class, and anxiety as sources of variation in school achievement. *J. Educ. Psychol.*, 53, 316-322.

Sarason, S. B. (1957). Test anxiety, general anxiety and intellectual performance. *J. Consult. Psychol.*, 21, 485-490.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R., and Ruebush, B. K. (1960). *Anxiety in Elementary School Children.* New York : Wiley.

Spence, K. W. (1958). A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *Amer. Psychologist*, 13, 131-141.

Stern, J. A., and McDonald, D. G. (1965). Physiological correlates of mental disease. *Ann. Rev. Psychol.*, 16, 225-264.

Welford, A. T. (1965). Stress and achievement. *Aust. J. Psychol.*, 17, 1-11.

(*Manuscript received* 23*rd March*, 1968)

TEACHERS' MARKS AND PUPILS' EXPECTATIONS: THE SHORT-TERM EFFECTS OF DISCREPANCIES UPON CLASSROOM PERFORMANCE IN SECONDARY SCHOOLS

BY ANTHONY J. PICKUP AND W. S. ANTHONY
(*School of Education, University of Leicester*)

SUMMARY. The research investigated (*a*) pupils' expected marks for a piece of work and their relationship to the actual mark for that work and (*b*) the effects of discrepancies between expected marks and teacher's marks. In classroom tests of normal classwork, it was found that actual marks tended to be less than expected marks, especially when actual mark was low. Discrepancies between pupils' expectations and teacher's mark were found to have no *single* motivational effect. However, there was a relationship between effectiveness of discrepancies and actual mark; the treatment of giving less marks than a pupil expected showed a slight tendency to be more effective than the opposite treatment when actual (objective) mark was high, but less effective when it was low. Thus it appears that the most usual treatment differed from the most effective treatment.

I.—INTRODUCTION.

WHEN a teacher asks his pupils to perform tasks for which they will be returned marks or grades he is initiating an activity which is one of the most extensively applied methods in use in education. The situation which arises when a mark or grade is returned to a pupil is at first sight a simple application of an essential 'feedback' process; a method by which the teacher can assess the success of his teaching and by which a learner can assess the adequacy of his learning. In the typical educational context however, the process of returning marks to pupils on completion of a task is not merely informational. The situation contains a number of elements which may affect the subsequent motivational state of the pupil. The simplest motivational element is the avoidance of the lack of interest which would result from no knowledge of results, but other elements are also present. A good mark may be considered a reward and a poor mark a punishment. Marks may be accompanied by statements or comments so that together or singly they constitute a situation of praise or of blame. The motivational state of the learner may be influenced by his perception of the likelihood of success or failure and the perceived status of the teacher may influence the reception of marks, grades or comments.

The now notorious unreliability of marks which may be awarded (Vernon 1962, Britton *et al.* 1966) may be further complicated by the teacher deliberately distorting marks for motivational purposes (Vernon 1962, p. 33). The practice of deliberately distorting marks introduces a dangerous assumption. The teacher who does this presumably makes a number of assumptions about the effects which a mark will have upon a pupil's later performance. Is a low mark (whatever that may mean) necessarily discouraging?

It is reasonable to assume that when a pupil has completed a task for which he is to receive a mark or grade, the pupil forms an expectation of some kind regarding this mark or grade. This expectation may be of a crude nature; a 'good' mark or a 'poor' mark. It may be further refined in relation to the pupil's past performance; a "better mark than usual" or a "worse mark than usual." If the teacher asks his pupil to state what he (the pupil) considers his work is worth, this evaluation may be the closest approximation to the pupils 'true' expectation that can be found. Any vague 'hopes' which may be entertained by the pupil will be modified by some realism, because the recipient of his evaluation is the teacher who will mark his work. This evaluation is therefore described throughout the present report as the pupil's expectation.

The aims of the research here described were to examine, *in typical classroom settings*, the following:

(*a*) The relationship between pupil's expectations and the teacher's mark.

(*b*) The effects upon subsequent performance of a discrepancy between the pupil's expectation and the mark awarded by the teacher.

II.—PREVIOUS RELATED WORK.

A search of the literature concerned with educational incentives revealed no previous work directly related to this topic. While a number of analogies may be drawn with research in the area of "praise or blame," these do not present precise guidelines for the formation of hypotheses. Kennedy and Willcutt (1964) in their review of literature in this area, have shown that the experimental history of praise and blame studies has failed to produce clear-cut conclusions, though recent evidence appears to support the belief that praise is a superior incentive to blame. "Level of aspiration" studies offer further analogies, limited by the fact that goal-setting is a pre-performance factor while the present research is concerned with the post-performance factor of expected mark. Child and Whiting (1949) illustrate a number of widely accepted generalisations in the area of level of aspiration. Murstein (1965), studying the relationship of grade expectations to grades actually received, found that students did not change their expectations of grades as a result of receiving actual grades. High ability students expected high grades and received them. Low ability students similarly expected high grades and did not modify their optimism as a result of experience.

The authors were impressed with the method of the Page (1958) research, concerned with the effectiveness of teachers' comments. This research has been popular with reviewers (Campbell and Stanley 1963, Charters and Gage 1963) because of its avoidance of many pitfalls in a 'generalisable' design. The experimental methods used in the research described below are evolved from the Page (1958) research.

III.—PLAN OF INVESTIGATION.

Preliminary Work.

A six-class study, described in full elsewhere (Pickup 1967, ch. 6), was carried out to examine the relationship between teacher's marks and pupil's expectations which would occur in schools under non-experimental conditions. Measures were taken of pupils' expected marks following classroom tests and of teachers' marks for the tests. Precautions taken to avoid the 'artificiality' of experiment included the use of the usual classroom teacher as a sub-experimenter. The main findings of this study were as follows:

(1) Out of a final sample of 123 pupils three had estimated their mark correctly, relative to the teacher's mark.

(2) Of those pupils who were not accurate in their estimate 68·3 per cent overestimated their mark and 31·7 per cent under-esimated their mark on the first occasion of being asked to evaluate their work. (This is a highly significant difference, $P < \cdot 001$.)

(3) Girls who had over-estimated their mark on the first occasion modified this optimism significantly more than over-estimating boys when both were asked to evaluate their work on a second occasion.

(4) Low pre-test scorers over-estimated more often than high pre-test scorers. (This was a highly significant difference, $P < \cdot 001$.)

Result (4) listed above, the correlation between pre-test score and frequency of over-estimation, arose in the following way: The data were examined for indications of a possible effect of discrepancy between expected and teacher's mark upon subsequent test performance. It was recognised that *if* there were a correlation between pre-test score and discrepancy, then the well-known phenomenon of statistical regression (Campbell and Stanley 1963, p. 181) would produce the false appearance of an effect of discrepancy upon subsequent performance. Therefore result (4) was looked for, and the actual finding of the result meant that this particular study would not yield a true picture of the possible effect of discrepancy. In any case, this was a correlational study and hence less satisfactory than an experimental study in elucidating cause-and-effect relationships. It was therefore decided to carry out an experimental study in which the discrepancy between expectation and test-score would be manipulated as an independent variable, while retaining (as far as possible) the real-life aspects of the situation.

A two-class pilot study was carried out to test a possible experimental design. This study (Pickup 1967, ch. 7) yielded useful information which led to the improved design described below.

The Main Study.

The major purpose of the main study was to examine by experimental means the effects upon subsequent classroom performance of a discrepancy between a teacher's mark and a pupil's expectation. A secondary purpose was to check the findings of the preliminary work described above.

Six non-selective single-sex secondary schools provided a final sample of 253 pupils from ten classes. The school subjects used in the experiment were English and Mathematics.

IV.—THE MEASURES USED.

Performance Tests.

Participating teachers were requested to set two tests of their normal classwork. The material to be tested was to be as independent of previous learning and experience as possible. The teachers were asked to construct their tests as follows:

(*a*) A section of 20 ' objective ' questions which could be scored easily on a right/wrong basis.

(*b*) A ' woolly ' question containing a subjective element. The pupil should not be able to assess easily whether marks are gained or lost on this question. (This question was never marked because its sole puprose was to disugise the fact that marks had been tampered with.)

(*c*) The tests were to be so constructed as to give every pupil an opportunity of gaining some marks, but extreme difficulty in obtaining full marks.

These tests were to be marked by the experimenter, following a marking scheme devised by the participating teacher. Teachers were informed that the marks returned after the first test may differ from the actual performance of the pupil.

Pupils' Expected Marks.

Pupils were asked to enter upon an additional slip of paper the mark (out of 20) that they believed their test paper to be worth. These assessment slips were to be given to the experimenter together with the test scripts.

Additional Information.

The Junior Eysenck Personality Inventory (Eysenck 1965a, b) was administered to the pupils during a testing session separate from the experiment A measure of the pupils' academic motivation was obtained during the same session by means of a modified 19-item test described by Buxton (1965). A measure of the participating teachers' leadership attitude was taken by means of a modified version of the Assumed Similarity of Opposites test described by Fiedler (1958).

V.—PROCEDURE.

Participating teachers were supplied with an instructional leaflet concerning their part in the experiment. After the administration of the pre-test and collection of the pupils' assessment slips, the experimenter collected this material and marked the pre-test scripts following the teacher's marking scheme; these marks are referred to as ' actual ' marks. All pupils who had indicated that they had been absent for some portion of the work tested in the pre-test were assigned to a ' No Treatment ' group. Similarly, all pupils who had recorded an expected mark of 0–3 or 17–20 were assigned to the No Treatment group. (However, in one class of 26, 10 pupils expected a mark of 17–20, and in this case only those pupils who would have fallen into the High Treatment group were excluded, which resulted in a loss of three pupils).

Pupils remaining after these exclusions were randomly assigned to one of three treatment groups (High, Low and Matching treatments) within levels of ' actual ' score rank. Thus, within any one class-list, actual scores were converted into rank order ; the three top ranks (1, 2, 3) were grouped into a triplet, the next three ranks (4, 5, 6) were grouped into a next triplet, and so on. Ties had previously been ranked by the toss of a coin. Two numbers on an ordinary six-sided die were selected to represent each of three treatments, for the purpose of random assignment : within each triplet, the first member was assigned to a treatment by a throw of the die ; the second member was similarly assigned to a second treatment · and the third member was assigned to the remaining treatment. This procedure was repeated for each triplet.

The experimenter administered experimental treatments by writing, on a mark list, a mark which was 3 marks above (' High ' treatment), 3 marks below (' Low ' treatment), or exactly matching (' Matching ' treatment) the pupil's expected mark, according to the treatment group in which that pupil fell. This mark list was returned, with the pupils' assessment slips, to the participating teacher. The teacher entered the mark from this list on to each pupil's assessment slip and this was returned to the pupil (i) immediately prior to the work which was to be tested on the post-test or (ii) immediately prior to the post-test where (i) was not possible. (An excuse was made to the pupils that the actual test scripts were not available at that moment ; in fact neither pre-test nor post-test scripts were ever returned to the pupils.)

Post-test scripts, marking scheme and assessment slips were collected by the experimenter and a post-test mark list, based upon the teacher's marking scheme, was returned to the teacher with the assessment slips. During this visit to the school, the experimenter administered the personality and academic motivation tests to the pupils while the participating teacher completed the ASO test in another room. All pupils were eventually informed that an experiment had taken place and that the scripts were required for records.

The percentile rank of each raw score in a class test was calculated and converted into a T score (Garrett 1958, pp. 314-318, Table G). Thus each set of test scores has a normal distribution with a mean of 50 and a standard deviation of 10.

VI.—RESULTS.

A.—*Pupils' expected mark and actual mark.*

A1. When the pre-test data were examined it was found that 72·6 per cent of the pupils over-estimated their mark relative to the actual mark. A considerable number of pupils (almost 10 per cent)were accurate in their estimate. Of those who were not accurate in their estimate, 81·3 per cent over-estimated their mark. This tendency to over-estimate is highly significant ($P < \cdot 001$) and confirms the finding of the preliminary work.

A.2. The data were examined for significant direction of change of expectation following experimental treatment. It was found that girls who been exposed to Low treatment following the pre-test showed a highly significant tendency ($P < \cdot 001$) to reduce their expectations for the post-test. No other significant directions of change were found.

A.3. Pre-test score was correlated with the difference between actual mark and expected mark ($r = 0 \cdot 33$, $P < \cdot 001$). That is, low test scorers were especially liable to over-estimate their marks, as had also been found in the preliminary work.

B.—*The effects of the experimental treatment.*

B.1. The effects of each treatment upon gain scores were examined by means of the analysis of variance by ranks. Each intact triplet in the experiment yielded ranks for a Friedman two-way analysis of variance (Siegel 1954, pp. 166-172). The ranked data from each triplet were cast into various categories to examine treatment effects. There were no significant effects here. Other analyses, carried out to examine possible subject differences, also proved non-significant.

B.2. Analyses concerned with the exploration of treatment effects in relation to personality factors, academic motivation and teachers leadership attitude, proved non-significant except for some minor sex-differences.

B.3. Examination of the data on High and Low treatment and pre-test position in class showed a correlation between pre-test score and differential effect of treatment. It is to be noted that in this experimental design, treatments are applied with equal frequency within each level of pre-test score, and therefore there is no misleading effect of statistical regression as there was in the earlier correlational study. This correlation was at first tested by relatively crude techniques which failed to yield significant results, and then by the product-moment method, as follows. The variables studied were matched-pairs difference scores and mean pre-test scores. Each randomised triplet in the experiment was examined for losses of individuals ; any triplet which had lost a high or low treatment pupil was excluded from analysis. (i) The standardised scores for the high treatment pupil and the low treatment pupil yielded a mean pre-test score for that pair. (ii) The gain score for the high treatment pupil was subtracted from the gain score for the low treatment pupil to yield a difference score for that pair. A product-moment correlation between the variables in (i) and (ii) yielded r=0·275 (N=74 pairs). The significance of this obtained r was tested (Walker and Lev 1953, p. 251) yielding t=2·43, df.=72· This correlation is significant beyond the ·02 level (two-tailed test).

This significant correlation indicated the need to examine all differences between treatments in the same way. No significant differences were found other than between High and Low treatment.

A crude indication of treatment effects related to pre-test score is to be found in Figure 1. In this graph three performance groups (pre-test) 'failing' pupils (45 marks and less) 'average' pupils (between 46-54 marks) and 'successful' pupils (55 marks and more) are studied to show average gain scores following experimental treatment. The general downward slope is a reflection of statistical regression, but the significant aspect of the figure is the difference between experimental treatments at the various levels of pre-test score.

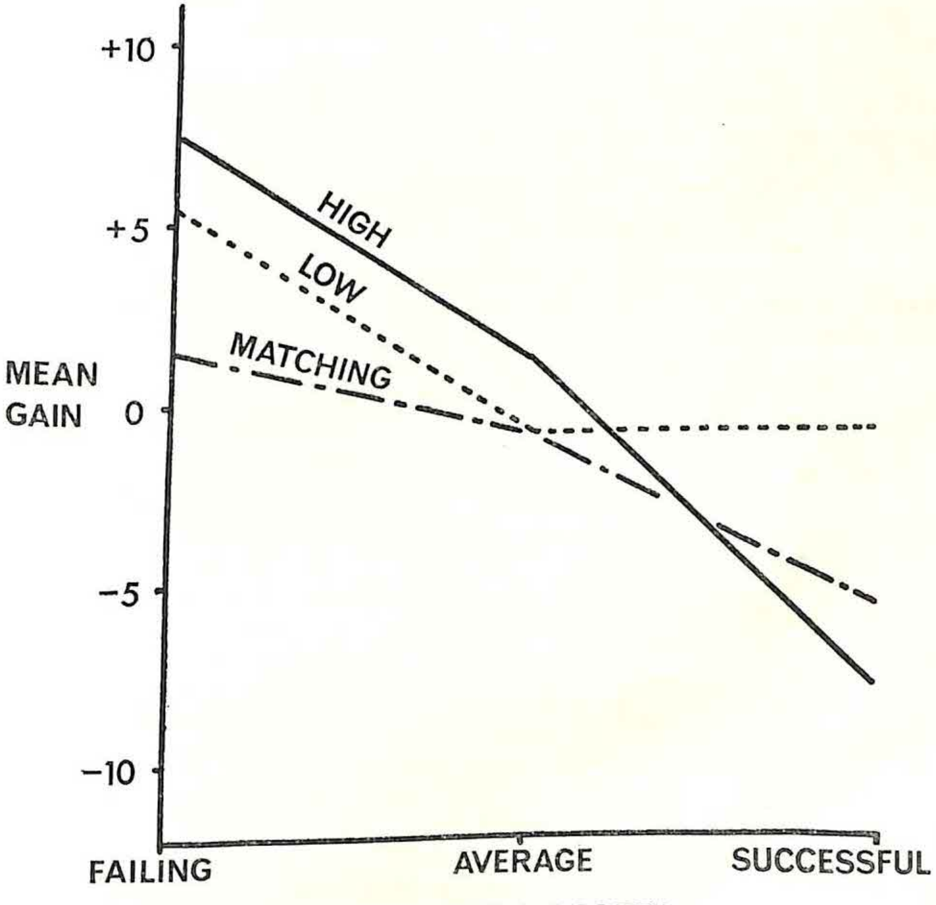


FIGURE 1

THE EFFECTS OF THE THREE EXPERIMENTAL TREATMENTS AT DIFFERENT LEVELS OF PRE-TEST SCORE. THE GAIN SCORE IS THE POST-TEST T SCORE MINUS THE PRE-TEST T SCORE.

VII.—CONCLUSIONS AND DISCUSSION.

This research aimed to investigate (*a*) the relationship between pupils' expected marks and actual marks and (*b*) the effects, upon subsequent performance, of discrepancies between expected marks and marks received by the pupil from the teacher. The method of the research was designed for generalizability to the normal school setting; the only divergence from normal classroom procedure was in the collection of pupils' expected marks. It was intended, and it is believed that pupils were unaware that observation or experiment were taking place.

The results naturally fall into two categories corresponding to the two aims of investigation.

(*a*) Concerning the relationship between expected and actual marks, it was found that (1) a large majority of pupils over-estimated their mark on the first test; (2) girls who expected more marks than they received showed a tendency to reduce their expectation on the second test; (3) low test scorers were especially liable to over-estimate their mark.

(*b*) Concerning the effects upon performance of discrepancies between expected and received mark, it was found that (1) there was no significant overall effect of discrepancies; (2) there was no significant effect of discrepancies when school subject, sex of pupil, personality of pupil or teacher's leadership attitudes was taken into account, except for some minor sex-differences; (3) effect of discrepancy was significantly related to pre-test score: giving a pupil more (rather than less) marks than he expected was found to be the more effective treatment of the 'poorer' pupils but the less effective treatment of the 'better' pupils.

Results (*a*) (3) and (*b*) (3) should be considered together. Result (*a*) (3) and the preliminary work, indicated that low test scorers especially score *less* than they expect. Result (*b*) (3) suggests that it would be a more effective treatment to give low test scorers *more* than they expect; high scorers seem to perform better if they have been given less than they expect. So far as the low scorers at least are concerned, it appears that the better treatment is in the opposite direction to what tends to occur. Any beneficial effect of a novel treatment may, of course, be due only to its novelty.

Implications for further research.

From the present study it seems that if there is a single motivational effect of a discrepancy between a pupil's expected mark and the mark received from the teacher, this effect is so small as to be insigificant in the normal classroom setting. Such an effect might, however, appear with the use of more sophisticated performance tests and laboratory-type conditions. Other experiments employing a method similar to that employed in the main study could investigate the effects of repeated discrepancies over a longer period of time. Some participating teachers suggested that the procedure of asking a pupil for a written assessment of his own work was in itself a motivating influence; this is worth investigating.

It seems that the effectiveness of one treatment relative to its opposite, as an incentive, reverses as the opposing treatments are applied at different levels of test-score in the class. This 'reversal effect' could conceal a treatment effect if positive effects of a treatment at top ability levels are 'cancelled-out' at the bottom levels. In no study of educational incentive which has been found has the possibility of this reversal effect been considered. Further investigation of the effect would be desirable.

ACKNOWLEDGEMENTS.—Thanks are due to the Chief Education Officers of the London Boroughs of Croydon and Bromley and to the head teachers and participating staff of the ten schools in which the research took place. Thanks also are due to Dr. Sybil B. G. Eysenck, who kindly supplied the Personality Inventories used.

VIII.—REFERENCES.

BRITTON, J. N., MARTIN, N. C., and ROSEN, H. (1966). *Multiple Marking of English Compositions.* Examinations Bulletin No. 12. Schools Council. H.M.S.O.

BUXTON, C. E. (1965). Evaluation of forced-choice and Likert-type tests of motivation to academic achievement. *Brit. J. Educ. Psychol.*, 36, 192-201.

CAMPBELL, D. J., and STANLEY, J. C. (1963). Experimental and Quasi-experimental designs for research on teaching. In N. L. Gage (Ed.) *Handbook of Research on Teaching.* Chicago : Rand McNally.

CHARTERS, W. W., and GAGE, N. L. (1963). *Readings in the Social Psychology of Education* Boston : Allyn and Bacon.

CHILD, I. L., and WHITING, J. W. M. (1949). Determinants of level of aspiration. *J. Abn. Soc. Psychol.*, 44, 303-314.

EYSENCK, S. B. G. (1965a). *Junior Eysenck Personality Inventory.* London : Univ. Lond. Press.

EYSENCK, S. B. G. (1965b). *Manual of the Junior Eysenck Personality Inventory.* London : Univ. Lond. Press.

FIEDLER, F. (1958). *Leader Attitudes and Group Effectiveness.* Urbana : Univ. Ill. Press.

GARRETT, H. E. (1958). *Statistics in Psychology and Education* (5th Edn.). London : Longmans Green.

KENNEDY, W. A., and WILLCUTT, H. C. (1964). Praise and blame as incentives. *Psychol. Bull*, 62, 323-332.

MURSTEIN, B. I. (1965). The relationship of grade expectations and grades believed to be deserved to actual grades received. *J. Exp. Educ.*, 33, 357-362.

PAGE, E. B. (1958). Teacher comments and student performance. *J. Educ. Psychol.*, 49, 173-181.

PICKUP, A. J. (1967). *Teachers' Marks and Pupils' Expectations : the Short-term Effects of Discrepancies Upon Classroom Performance.* Unpub. M.Ed. Thesis. Univ. of Leicester.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* New York : McGraw-Hill. Int. Stud. Edn.

VERNON, P. E. (1962). *The Measurement of Abilities.* London : Univ. Lond. Press.

WALKER, H. M., and LEV, J. (1953). *Statistical Inference.* New York : Holt.

(*Manuscript received* 28*th March*, 1968)

RESEARCH NOTES

A PROBLEM CHECK-LIST FOR TEACHERS

By SEYMOUR LEMESHOW

(*Educational Development Clinic, Jersey City State College, N.J.*)

Summary. The *Teacher Operational Problems-Identification* is a self-administering check-list of 445 problems faced by teachers in eight main areas of their work: persons in authority, peers and staff, the community, students and discipline, working conditions, methods and curriculum, personal and social, and professional role. The instrument is intended to help teachers identify and confront their problems as a first step towards solving them, and can be used in the counselling of teachers, in-service education and research. The list of problems was validated by 500 secondary classroom teachers, each item requiring a minimum of five responses for inclusion. The coefficient of reliability based on a retest of 25 teachers was ·96. Suggestions are made for further research and the development of norms.

I.—Rationale and purpose.

Teachers inevitably encounter a variety of problems arising from their work in the classroom and from their relations with colleagues, administrators, parents and the community at large. On the assumption that many teachers are unable to identify readily the nature of their problems, the author (Lemeshow, 1966a, 1966b) has constructed a self-reporting, self-administering check-list to help them take the first step towards finding a solution. But the check-list, the *Teacher Operational Problems-Identification*, is designed to be more than an aid to self-knowledge and evaluation. By drawing attention to those specific areas of difficulty the individual teacher is ready to talk about, the list can guide those from whom he seeks advice. The check-list could be used in in-service education for getting problems out into the open, so helping the group to work more freely and with less inhibition towards their solution. Finally, research into the teaching situation may be facilitated by this method of assessing the nature of teachers' problems. A number of variables which may be related to teacher difficulties could be investigated and evaluated.

II.—Development of the instrument.

(i) *Discovery of Problems.* The development of the check-list employed techniques which were analogous to, but peculiarly different from, existing instruments which are essentially directed at student groups in school settings. The sources of the problems used by Mooney and Gordon (1950), Billett and Starr (1958) and Remmers and Shimberg (1949) were mainly based on the analysis of essays written by students. In the present study, however, it was assumed that qualified educators could furnish statements of problems directly when provided with an appropriate questionnaire. Accordingly, a panel of 96 contributors representing 25 states and composed of 11 classroom teachers, 34 school administrators, 29 teacher-educators, 6 guidance personnel and 16 other specialists submitted statements of problems which came to their awareness as a result of their experiences.

A wide and varied list of 1,714 problem statements was contributed by the panel. These problems were screened, edited, grouped and regrouped until 720 distinct problems were identified. A panel of nine teachers then read the statements and offered suggestions for language clarification.

(ii) *Validation of Problems.* A questionnaire containing the 720 refined problem statements was distributed to 500 secondary teachers covering the whole range of the curriculum and representing 28 states. The teachers were asked to identify the statements on the questionnaire which they recognised as being their problems.

Only those problems identified by at least 1 per cent of the validating sample (that is, by five teachers) were eligible for the instrument. In fact, many of the problems were identified by from 10 to 49 per cent.

(iii) *Format of the Check-List.* The 445 problem statements which met the minimum requirements for validation were then grouped into eight categories determined by a logical analysis of the statements. A panel of three judges judged the appropriateness of the categories and sorted the problem statements into areas of best fit. Approximately 67 per cent of the problems were judged unanimously for category assessment, while 33 per cent were agreed upon by two of the three judges. The problem area designations, the number of items subsumed within each area. and examples of the most frequently encountered problems within each area are given in the Appendix. Opportunity is provided on the checklist for the respondent to indicate the frequency and intensity of occurrence of the problems pertaining to him.

III.—RELIABILITY OF THE INSTRUMENT.

The problems experienced by individuals change. This instrument is designed to reflect such changes. The teacher will reveal only what he chooses or what he can reveal at the moment. An instrument designed to assess the ephemeral nature of problems would not be one which would be amenable to the usual standards of reliability.

The *Teacher Operational Problems-Identification* is not a test or a measuring device. It provides no precise measurement characteristics. It would be difficult, therefore, to employ measures of reliability such as the split-halves or the alternate form method. However, since one of the uses of this instrument is for survey, its coefficient of stability was of interest.

The reliability sample consisted of 25 teachers, median age 25·5 years, who responded to the check-list on two occasions a week apart. The coefficient of stability was found to be 0·96, which suggests that the instrument is sufficiently stable for such a purpose as a survey of problem areas for in-service planning.

IV.—NORMS.

Since the *Teacher Operational Problems-Identification* provides for a self-identification of the problems the teacher himself perceives, there can be no scores in the sense of precise measurements. There are instead item and area counts. However, there may be a need to make comparisons between the frequency of problems an individual faces and the extent to which such problems occur in the population to which he belongs. The next step in this investigation will be to explore the feasibility of developing sectional and regional norms which would provide for a meaningful interpretation of the problems of individuals or groups. In the meantime it is suggested that individual users should establish local norms for their own purposes.

V.—REFERENCES.

BILLETT, R. O., and STARR, I. S. (1958). *Manual for Billett-Starr Youth Problems Inventory.* Yonkers-on-the-Hudson : World Book Co.

LEMESHOW, S. (1966a). *The Development of a Self-Reporting, Self-Administering Check List of Teacher Operational Problems.* Unpublished Ed.D. final document. New York University.

LEMESHOW, S. (1966b). *Teacher Operational Problems-Identification.* Princeton, New Jersey : Co-operative Test Division, Educational Testing Service (Research Edit.).

MOONEY, R. L. and GORDON, L. V. (1950). *Manual for Mooney Problems Check Lists.* New York : Psychological Corporation.

REMMERS, H. H., and SHIMBERG, B. (1949). *Manual for SRA Youth Problems Inventory.* Chicago : Science Research Associates.

(*Manuscript received 6th January,* 1968)

VI.—APPENDIX.

The problem areas and examples of the problems most frequently encountered within each area by the 500 teachers in the validating sample.

Area	Description of Area and Examples of most frequently encountered problems.	Total number of teachers responding out of 500	
			%
"A"	PERSONS IN AUTHORITY (60 items)		
	Tending to feel nervous when observed	78	(15·6)
	Administrators being unduly influenced by public	73	(14·6)
	Not agreeing with administrators on certain policy	65	(13·0)
	Not being involved in decision making	52	(10·4)
"P"	PEERS AND STAFF (30 items).		
	Former teachers preparing my students poorly	68	(13·6)
	Bothered by teachers not understanding students	65	(13·0)
	The teacher next to me unable to control his class	63	(12·6)
	Colleagues not following the school policy	51	(10·2)
"C"	THE COMMUNITY (35 items)		
	Parents lacking concern about failures	137	(27·4)
	Parents lacking interest in students' progress	81	(16·2)
	Unhappy with the P.T.A.	64	(12·8)
	Little interest by community in academic work	47	(9·4)
"S"	STUDENTS AND DISCIPLINE (65 items)		
	Concerned about students' study habits	158	(31·6)
	Teaching those who do not want to learn	145	(29·0)
	Students lacking concern about failures	117	(23·4)
	Students being too grade conscious	106	(21·2)
"W"	WORKING CONDITIONS (70 items)		
	Having too many papers to correct	149	(29·8)
	Classes being too large	141	(28·2)
	Not having the time to help individual students	130	(26·0)
	Having too much clerical work	123	(24·6)
"M"	METHODS AND CURRICULUM (65 items)		
	Students being unable to read	219	(43·8)
	Needing realistic course for slow learners	137	(27·4)
	Needing to know how to challenge slow learners	117	(23·4)
	Needing materials to teach slow learners	106	(21·2)
"PS"	PERSONAL AND SOCIAL (60 items)		
	Feeling exhausted at the end of school day	131	(26·2)
	Needing to "moonlight" to make ends meet	86	(17·2)
	Needing Summer employment	72	(14·4)
	Being overweight	43	(8·6)
"PR"	PROFESSIONAL ROLE (60 items)		
	Teachers' salaries are too low	202	(40·4)
	Unable to keep up with my reading	173	(34·6)
	Wondering if I should stay in teaching	88	(17·6)
	Our profession not commanding respect	86	(17·2)

A NOTE ON THE STABILITY OF INDUCED CONSERVATION

By C. K. MACKAY AND T. M. KILKENNY
(*Department of Psychology, University of Aberdeen*)

SUMMARY. Inhelder, *et al.* (1966), imply that the progress made by young children when taught to conserve amount by the Frank-Bruner screening method is unstable. They suggest that a post-test after a two-week delay will show the specificity of the training procedure. In this study dealing with children from 4½–7 years of age, Frank's original screening experiment was replicated and two further post-tests were given two weeks later. The results of the screening experiment are similar to those of Frank. There is, however, no decrement in performance across the three post-tests.

I.—INTRODUCTION.

In a recent comment to the *American Psychologist*, Inhelder, Bovet, Sinclair and Smock (1966) suggest that the gains in children's conservation responses reported by Frank (in Bruner, 1964) were due to " the child's attention (being directed) to aspects of the task that gave him a momentary, economical solution to the problem." They imply that children who change from non-conservation to conservation responses during the screening procedure are being trained to conserve in a particular situation and that further exploration, e.g., a second post-test two weeks later, makes clear the specificity of this training. The present investigation attempted a replication of the original screening experiment (in detail in Bruner, *et al.*, 1966, Chapter 9). Two further post-tests, both administered two weeks after the training session, were added.

II.—METHOD.

The subjects, 77 Yorkshire nursery and infant school children, ranging in age from 4½–7 years, were subdivided into five age-groups as shown in Table 1. The procedure adopted with each age-group was that outlined by Bruner (1966).

The conventional pre-test employed has two parts. In both, two standard beakers are equated for level of coloured water. In part 1, one of the standard beakers is poured into a tall, narrow beaker and the child is asked whether there is now the same to drink in the two test beakers. In part 2, the liquid is poured from one of the standard beakers into six, small, narrow beakers and the child is again asked whether there is the same to drink in the remaining standard beaker and in the six smaller beakers taken together. The criterion for a conservation response is a correct response to both questions.

The three screening procedures with four pairs of beakers were then administered, viz. : (1) Screening, pouring, question, no unscreening ; (2) prediction of level, no pouring ; and (3) screening, pouring, prediction of level, unscreening and questioning. This was immediately followed by post-test 1—a repeat of the standard pre-test. The same criterion of correct responses on both parts of the test was used for assessing the presence or absence of conservation.

Two weeks later, this post-test was repeated (post-test 2). This was immediately followed by post-test 3, which employed a different medium, modelling clay. Two identical discs of clay were presented to the child. When these were accepted as containing the same amount of clay, the second disc was then cut into quadrants, the quadrants were piled on top of each other to form a block and the child was asked if there was now the same amount of clay in the two displays. In this test, there is obviously only one response to be scored.

III.—RESULTS.

The numbers in each group showing conservation, as defined above, are shown in Table 1. In parentheses are shown, as plus scores, the numbers progressing from non-conservation to conservation, and, as minus scores, those failing to maintain conservation, from the previous test.

TABLE 1

THE PATTERN OF ACCEPTABLE RESPONSE BY AGE GROUPS.

Age Group (in Yrs.)	N	Number of children showing conservation									
		Pre-test	Post-test I	Gains	Losses	Post-test II	Gains	Losses	Post-test III	Gains	Losses
4½–5	14	0	1	(+1)		3	(+2)		6	(+4)	(—1)
5–5½	16	1	3	(+2)		2		(—1)	3	(+1)	
5½–6	15	3	6	(+3)		5		(—1)	7	(+2)	
6–6½	16	11	13	(+2)		13	(0)		14	(+1)	
6½–7	16	5	10	(+5)		10	(0)		14	(+4)	
Totals	77	20	33	(+13—0)		33	(+2—2)		44	(+12—1)	

The figures in parentheses refer to the number of children gaining or losing from previous test.

The changes observed taking place between the pre-test and post-test 1 were similar to those reported by Bruner. In general, we found a significant increase in the number of children giving conservation responses between the pre-test and post-test 1. We support his observation that children over 5 years benefit more from this type of prompting than do those younger. Of particular interest was the finding that fewer children at the 6½–7 years age level than at the 6–6½ years level displayed conservation responses on the pre-test. Here we found, like Bruner, that children of the older age group are failing because of a misuse of the principles of compensation. Of the 11 in this group who were non-conservers on the pre-test, seven gave incorrect responses only to part 1 of the pre-test and these incorrect responses seemed to be due to the emerging ability to take account, albeit incorrectly, of the ratio of amount of liquid to size of container. Their failure is seen as due to the tendency to equate all larger magnitudes as being from one pool of 'bigness.' In particular, it meant saying that the level of liquid in the taller glass would be lower because the glass was 'taller' or 'bigger.' In the remaining four, three showed no sign of conservation in either part of the pre-test: in the 6–6½ years group, all five non-conservers failed both parts of the pre-test.

Our main interest, however, is in the stability of gains across the three post-tests. A McNemar Test for the significance of changes (Siegel, 1956) shows an improvement for total scores on post-test 3 over post-tests 1 and 2 significant at the ·01 level in both cases. There is no difference between scores on tests 1 and 2.

IV.—DISCUSSION AND CONCLUSIONS.

These results largely support Bruner's (1966) findings. They do not support Inhelder, *et al.* (1966) in their implication that the progress made by children in this training situation is transitory. In the over 6 age group, those children who progress to conservation from pre-test to post-test 1 maintain this progress across the two delayed post-tests. In the under sixes there is some instability in those individual children who do attain conservation but the degree of stability is much greater than would be expected were the training during the screening procedure providing merely 'a momentary, economical solution.'

The reason for the improvement in post-test 3 is not known. It may be that the task is easier in that it involves, to some extent, discontinuous materials, or since the material used is manipulated much more frequently by nursery and infant school children, conservation may be achieved, as Lovell and Ogilvie (1960) suggest, earlier with this medium than with liquids. Also, only one correct response is required on this particular test.

Finally, we recognise that this paper has contributed nothing to the resolution of the basic differences in theoretical approach between the Harvard and Genevan projects. We do conclude that progress in conservation, specified in the Frank-Bruner manner (Gruen, 1966), achieved by particular training methods, is more stable than Inhelder, *et al.*, imply.

ACKNOWLEDGEMENT.—The authors wish to thank the Headmistress, staff and pupils of Springbank Infant School, Earby, Yorks, for their co-operation and practical help.

V.—REFERENCES.

BRUNER, J. S. (1964). The course of cognitive development. *Amer. Psychologist*, 19, 1-16.

BRUNER, J. S., OLVER, R. R., and GREENFIELD, P. M., *et al.* (1966). *Studies in Cognitive Growth.* New York: Wiley.

GRUEN, G. E. (1966). Note on conservation: methodological and definitional considerations. *Child Developm.*, 37, 977-983.

INHELDER, B., BOVET, M., SINCLAIR, H., and SMOCK, C. D. (1966). On cognitive development. *Amer. Psychologist.* 21, 160-164.

LOVELL, K. and OGILVIE, E. (1960). A study of conservation of substance in the junior school child. *Brit. J. Educ. Psychol.*, 30, 109-118.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* Tokyo: McGraw-Hill.

(*Manuscript received 23rd February, 1968*)

COGNITIVE BIAS IN SCOTTISH FIRST-YEAR SCIENCE AND ARTS UNDERGRADUATES

BY C. K. MACKAY AND MARION B. CAMERON
(*Department of Psychology, University of Aberdeen*)

SUMMARY. This paper is a partial replication of Hudson's (1966) work with English schoolboys. It is intended to examine whether the relationship found in Hudson's sample between Arts-Science specialization and convergent-divergent bias will be shown by a sample of Scottish undergraduates who have not been subjected to the same degree of specialization in previous schooling. The results indicate that, where students have elected to specialize in their first year, the relationship between area of specialization and cognitive bias holds: when non-specialist Arts and Science students are compared such a relationship is not found.

I.—INTRODUCTION.

The relationship between cognitive bias and Arts-Science choice in school boys of high academic achievement has been intensively studied by Hudson (1966). He claims that divergers (defined as those who are relatively more successful in open-ended tests than in conventional intelligence tests) tend to specialize in Arts while convergers (those who do better on IQ tests than on open-ended tests) choose Science. Cropley and Field (1968) report similar findings in an investigation of the relationship between convergent intellectual bias and specialization in Science in Australian High Schools. These findings pose an important question for educationists, namely, whether cognitive bias determines Arts/Science specialization or whether the educational experience of following an Arts or a Science curriculum moulds or at least reinforces a particular style of thinking. Hudson prefers the former view. But, while regarding style of thinking as a function of personality largely formed in early childhood (Hudson 1966, p. 132) and while confirming the observation of Getzels and Jackson (1962) that convergers and divergers show considerably different personality patterns, he does not rule out the possibility that cognitive bias may be modified by later influences such as school experience.

A group similar to Hudson's 'clever schoolboys' in age and academic attainment but differing in the degree of specialization involved in previous school experience, may be drawn from first-year undergraduates in Scottish universities. The different patterns of secondary and university education in England and Wales and in Scotland are discussed fully in the Robbins Report (Robbins 1963, Part I, Vol. E): the principal differences relevant to this study are these. The equivalent of an English Sixth Form is not traditionally found in Scottish schools and pupils in Scotland go up to university a year earlier than do their English counterparts. University entrance requirements in Scotland discourage early specialization (especially among scientists) by requiring, as minimum qualification, passes at Higher Grade from two of the three groups, English, Mathematics—Science, Foreign Languages, and a pass at least at Ordinary Grade from the remaining group. The four-year Honours Degree course and the three-year Ordinary Degree course both allow, indeed demand, a wider area of study than is traditional in England.

Hence Scottish first-year undergraduates are not yet irrevocably categorised as Arts or Science specialists and indeed may choose to follow either of these courses or a more general course comprising both Science and Arts subjects after completing their first year of university studies.

The aim of this paper then, is to investigate the relationship between divergent and convergent bias and Arts or Science specialization in a sample of students who differ from the original 'clever schoolboys' sample mainly in the degree of previous subject specialization.

II.—Method.

Subjects were selected from the Ordinary Psychology class at Aberdeen University. As there is a strong tradition in this University of taking a course in Psychology as an 'outside' subject, this class, numbering over 400 students, provides a fairly broad sample of Arts undergraduates and a number of students from the Science Faculty. It is recognised that this is a restricted sample in the sense that extreme Arts and Science specialists may avoid this particular subject. The probable effects of this attenuation are discussed later. From this class, a sample was drawn to satisfy the following criteria. All subjects must be male, in their first year of university studies, aged under 20 years and must have attended a complete course of secondary education conforming to the traditional Scottish pattern. Thus restricted, the number of subjects available was 70. Of these, 23 were Science students and 47 were Arts. The Arts group could be further subdivided into 34 who were Arts specialists (i.e., studying Arts subjects exclusively) and 13 who were following a less specialized Arts course (e.g., including a Science subject in their courses). These two groups are referred to as Arts Specialists and Arts Intermediate, respectively.

Three tests were administered to the original group of 423 students. The Moray House Verbal Reasoning Test (Adult) I was followed at a different session by the two open-ended tests, "Uses of Objects" and "Meanings of Words" (Hudson 1966, Appendix B). The "Uses of Objects" test required the subjects to write down as many uses as they could think of for five common objects—e.g., a barrel. The second test required a list of meanings for ten common, but ambiguous words—e.g., port. The protocols of the selected group were then scored.

Raw scores on the Moray House Test were converted to linear standard scores. The open-ended tests were scored according to the total number of responses given, converted to standard scores and averaged ior each subject. On the basis of these two standard scores, differential scores were obtained by subtracting the averaged open-ended score from the Intelligence score. These differential scores ranged from +2·30 to −2·54. A positive differential score indicates relative superiority of Intelligence Test performance over open-ended test performance (convergent bias) and *vice-versa*. Such scores do not indicate absolute levels of performance on either test. Using Hudson's system, subjects were classified on these scores as Extreme Convergers, Convergers, All-rounders, Divergers and Extreme Divergers in the proportions 1 : 2 : 4 : 2 : 1. The 30 per cent at each extreme were then compared on Arts-Science choice.

III.—RESULTS.

The distribution of the selected 70 subjects is shown in Table1. The χ^2 associated with this distribution on the null hypothesis that Arts-Science choice is independent of cognitive bias is 5·62 (df=2, p> ·05). This result cannot be taken to support Hudson's findings.

TABLE 1

DISTRIBUTION OF SCIENCE AND ARTS STUDENTS (N=70) ON CONVERGENT-DIVERGENT BIAS.

	Convergers	All-rounders	Divergers
Science	7	13	3
Arts (specialist and intermediate)	14	15	18

TABLE 2

DISTRIBUTION OF SCIENCE STUDENTS AND ARTS SPECIALISTS (N=57) ON CONVERGENT-DIVERGENT BIAS.

	Convergers	All-rounders	Divergers
Science	7	15	1
Arts (specialists)	10	8	16

If, however, only Science and Arts Specialists are compared, i.e., by eliminating the Arts Intermediate group—the resulting distribution, shown in Table 2, yields a χ^2 of 14·59, df=2, p< ·001. This level of χ^2 is entirely due to the exceptional distribution of Divergers and All-rounders: the distribution of Convergers is exactly that predicted from the null hypothesis. These results strongly support Hudson's claim that divergent bias is associated with specialisation in Arts: there is no suggestion, however, that the Science students in this sample tend to convergent bias more than do the Arts specialists.

IV.—DISCUSSION.

A major difficulty encountered in this investigation was that of drawing a sample representative of first-year male students who had experienced the traditional pattern of Scottish senior-secondary schooling. Any attempt to contact all such students or a representative sample of them seemed likely to result in bias of a type which would not be readily assessable. Students who agreed to a postal request to undergo appropriate testing in standard conditions would be self-selected and would not, in the opinion of the experimenters, constitute a random sample of the population being investigated. The first-year Psychology class was a large, 'captive' group containing a cross-section of Arts students and a number of Science students. The nature of bias in this sample was considered to be more easily understood. It is a fair assumption that a broad sample of Arts undergraduates is found in the Psychology class but it is not so for Science students. The majority of first-year Science students do not include Psychology in their curricula. Indeed one would expect, from the previously published work, that the Science students in the present sample would be less convergently biased than subjects drawn at random from the Science faculty: extreme convergers among physical scientists would be those least likely to venture from home-ground to the no-man's land of Psychology. Hence the nature of the bias in this sample is considered to be in favour of the Arts group and to the exclusion of physical Science specialists.

This consideration makes the finding of a relationship between divergent bias and Arts choice in the Aberdeen sample even more striking. Arts specialists are biased towards divergent modes of thinking even when compared with those Science students who include at least one 'outside' subject in their degree courses. Table 2 shows that, while the Science students in this sample are no more convergently biased than would be expected under the null hypothesis of independence of cognitive bias and Arts/Science choice, the distribution of divergent bias is strongly in favour of the Arts specialists.

The inclusion of the Arts Intermediate group destroys this distinction between Arts and Science students. The inference is that this particular group is less divergently biased than the Arts specialist group. It also suggests that the earlier assumption that a greater degree of convergent bias would be found in more specialist physical scientists is justified. If so, one may return to Hudson's speculation (1966, pp. 94, 95) on whether cognitive bias may be reinforced by specialisation. In an educational system where specialisation is encouraged, Hudson finds a distinct relationship between cognitive bias and Arts/Science specialisation. This distinction does not arise in the present investigation where the subjects are not involved in a clear-cut choice between Science and Arts. It might be that, in an educational system where specialisation is encouraged, those who have little 'natural' bias to convergence or divergence, having elected a particular course of study, are influenced by the modes of thinking demonstrated by their tutors as appropriate to the discipline they have chosen. A system of secondary and further education based on a clear-cut choice between Science and Arts would then tend to produce a distinct differentiation in cognitive bias between the specialist groups. That this distinction does not arise when less highly-specialised undergraduates are compared may be evidence supporting the hypothesis that educational experience is a reinforcing or modifying factor in the development of particular intellectual styles.

V.—REFERENCES.

CROPLEY, A. J., and FIELD, T. W. (1968). Intellectual style and high school science. *Nature*, 217, No. 5135, 1211–1212.

GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence*. New York: Wiley.

HUDSON, L. (1966). *Contrary Imaginations*. London: Methuen.

ROBBINS (1963). *Higher Education. Report*. London: HMSO.

(*Manuscript received 1st May*, 1968)

AN ALTERNATIVE TO WARD'S FACTOR ANALYSIS OF WALLACH AND KOGAN'S 'CREATIVITY' CORRELATIONS

By FRANCIS FEE

(*Antrim Education Committee*)

Summary. The intercorrelations of Wallach and Kogan's tests of creativity and intelligence were factor analysed by the Multiple Group Factor method after factor identification by Centroid analysis of each set of tests. Four first order factors and two second order factors were derived. The second order factors differentiated clearly between the creativity and the intelligence tests although they themselves were slightly correlated. The four first-order factors could be described respectively as creativity-verbal, creativity-visual, non-verbal intelligence and attainment factors. These factors were compared with those found by Ward (**1967**).

I.—Introduction.

Ward (1967) has presented a factor analysis of the inter-correlations of Wallach and Kogan's (1965) tests of creativity and general intelligence using the Varimax and Promax methods of rotation following a Principal Components Analysis. Despite the use of these methods Ward's findings are not particularly clear in the simple structure sense and some of his factors are difficult to interpret, although Wallach and Kogan's conclusions are generally supported. The analysis described here was carried out sometime before Ward's results were available and it is based on a different approach. The factors obtained seem to be clearer than Ward's and may serve as a better description of the data.

The aim of the investigation was to examine in factor analytic terms Wallach and Kogan's contention that their data represented distinct areas of cognitive functioning, namely 'creativity' and 'intelligence.' The first step was to examine the unidimensionality or otherwise of each area, the information thus obtained being combined with the overall hypothesis of distinct cognitive areas to produce a simple structure matrix and, possibly, second order factors.

II.—Results.

With unity in the diagonals the centroid method was used to analyse Wallach and Kogan's tables of 'creativity' and 'intelligence' test intercorrelations (Tables 7 and 8) separately. For each set of intercorrelations two factors were found with variance greater than unity, no significant third factor being indicated in either case. The second, bipolar factor and a graph of both factors derived from the 'creativity' intercorrelations suggested an oblique solution through two clusters of tests, one cluster comprising what Wallach and Kogan refer to as 'verbal' tests, with the exception of Instances—number, the other comprising the 'visual' tests with the addition of Instances—number. The centroid factors derived from the 'intelligence' tests were less clear. The STEP and SCAT tests with the exception of STEP mathematics formed a distinct cluster contrasting with the WISC picture arrangement and block design subtests. WISC vocabulary and STEP mathematics lay between the other STEP and WISC tests.

On the basis of these findings an hypothesis matrix was formed and applied to the total matrix of test intercorrelations (Tables 7, 8 and 9), again with unity in the diagonals. The Multiple Group Factor method as described by Horst (1965) and the method he describes for deriving second order factors and their transfer to the first order realm were used to produce the data in Table 1. Factors A and B, the first order representations of the second order factors, were rotated to simple structure on the basis of an hypothesis matrix contrasting the 'creativity' tests with the 'intelligence' tests.

TABLE 1

SUPERMATRIX OF FIRST-ORDER GENERAL FACTOR MATRIX AND FIRST-ORDER RESCALED SIMPLE STRUCTURE FACTORS OF TWENTY COGNITIVE TESTS (WALLACH AND KOGAN, 1965).

Test	Factor					
	A	B	I	II	III	IV
1. Instances-uniqueness	420	—045	531	—072	036	—152
2. Instances-number	311	053	—044	282	068	000
3. Alternate Uses-uniqueness	639	—023	652	011	—019	—056
4. Alternate Uses-number	592	047	507	084	—021	038
5. Similarities-uniqueness	593	—029	668	—038	—012	—079
6. Similarities-number	641	053	593	058	—022	035
7. Pattern meanings-uniqueness	407	004	050	287	—037	081
8. Pattern meanings-number	298	—078	—077	296	011	—051
9. Line meanings-uniqueness	430	108	072	288	036	088
10. Line meanings-number	481	—018	087	318	—013	027
11. WISC vocabulary	050	483	116	—048	076	371
12. WISC picture arrangement	064	333	028	026	394	—189
13. WISC block design	—051	470	—086	022	360	006
14. SCAT-verbal	000	706	023	—009	—024	742
15. SCAT-quantitative	—052	692	—093	033	—012	727
16. STEP-mathematics	033	667	116	—060	087	540
17. STEP-science	001	703	—018	017	028	676
18. STEP-social studies	013	729	035	—013	024	700
19. STEP-reading	—032	693	—026	—002	—052	772
20. STEP-writing	—039	706	—078	032	—054	796
Intercorrelations	1·000		1·000			
	—235	1·000	—531	1·000		
			—072	020	1·000	
			—134	025	—501	1·000

III.—DISCUSSION.

The factors A and B in Table 1 appear to confirm Wallach and Kogan's conclusion that they had succeeded in defining "a dimension of individual differences . . . which, on the one hand, possessed generality and pervasiveness, but which, on the other hand, nevertheless was quite independent of the traditional notion of general intelligence," although the fact that these two factors are slightly correlated would seem to modify the claim that the dimensions are ' quite independent " of each other. Factors I to IV bear out the hypothesis matrix to a very large extent there being very few values in the zero positions exceeding ±0·1. This matrix seems to be very much ' cleaner ' than that presented by Ward.

Despite the seemingly odd behaviour of the Instances—number test, Factors I and II can be described as ' creativity-verbal ' and ' creativity-visual," respectively. The small but positive loading of the WISC vocabulary subtests supports the verbal nature of Factor I and suggests that it might be of more interest to examine tests of creativity in the context of tests representing the various factors of ability rather than the crude estimates of global or general intelligence used by Wallach and Kogan.

Factor III can be regarded as a factor of mainly non-verbal intelligence while Factor IV is clearly an attainment factor. There is no very obvious reason why Instances—uniqueness and WISC picture arrangement have small but marked negative loadings on Factor IV.

Ward's Factors II and III correspond to Factors I and II of the present analysis although they have been rotated differently, his Factor II appearing as a general factor of creativity and his Factor III being poorly defined. Ward suggests that his Factor III reflects " scoring for number of responses as opposed to uniqueness or accuracy of responses." Despite the loadings of Instances—number the ' creativity ' factors presented here seem rather more meaningful. Ward's Factors IV and I cover the same cognitive area as Factors III and IV of this analysis, but Ward's rotation gives his Factor IV the appearance of a general factor somewhat difficult to explain away. Again the factors presented here seem to be more meaningful.

IV.—CONCLUSIONS.

The above analysis supports Wallach and Kogan's view that they have established a ' creativity ' dimension relatively independent of general ability as measured by the usual tests of attainment and intelligence although this independence may not be as complete as Wallach and Kogan maintain. ' Creativity ' is clearly not unidimensional and there is a hint that it may have relationships with factors of ability not explicitly taken into consideration by Wallach and Kogan and not specifically represented in their set of tests.

Although a factor-analytic approach different to that of Ward's was used it is not clear why the factors shown here should differ so much from his.

V.—REFERENCES.

HORST, P. (1965). *Factor Analysis of Data Matrices.* New York : Holt, Rinehart and Winston.

WARD, J. (1967). An oblique factorization of Wallach and Kogan's ' creativity ' correlations. *Brit. J. Educ. Psychol.*, 37, 380-382.

WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children : A Study of the Creativity–Intelligence Distinction.* New York : Holt, Rinehart and Winston.

(*Manuscript received 1st April, 1968*)

COMMENT

BY JAMES WARD

It would appear that in these analyses of the Wallach and Kogan correlations two distinct aims were involved : Fee sought to " examine in factor analytic terms " an important claim made by the authors, whereas I was attempting to obtain " factorial data which would aid in the further evaluation of particular tests." Even so, though different methods were adopted, both sets of results suggest that a separate dimension of creativity has been established and we thus agree substantially on the major issue. However, there are a number of points of procedure and interpretation which merit comment.

In the first place, centroid methods of analysis are at best an approximation to principal component analysis and may well give rise to results at variance with the more mathematically rigorous procedure. Secondly, subjective approaches such as are involved in the construction of a hypothesis matrix are clearly inferior to objective methods based on analytic rotation and which do not prejudge the final result. It is true that both the Horst method and Promax are aimed at simple structure, but the achievement of a ' clean ' solution does not imply automatically that it is best either in the mathematical or interpretive sense. Thus Fee does not seem to have analysed the complete set of intercorrelations before arriving at his hypothesis matrix and this may account for the disparities between the two solutions. Having analysed this data several times I find his creativity factors I and II somewhat difficult to accept, particularly in the light of the comparatively high negative correlation quoted, which when considered against the original correlations is rather implausible. Factor IV poses no problems but Factor III certainly does not deserve identification as a factor of non-verbal intelligence. The loadings on the two WISC performance tests are moderate and in any event non-verbal tests are not adequately represented in the battery. Whilst acknowledging some interpretative difficulties over Factors II and IV in my own analysis. I must question therefore Fee's statement that the factors he presents seem to be more meaningful.

I concur however with Fee's observation that creativity tests are better examined in the context of multifactorial test batteries in which the various factors of ability are sampled, together presumably with personality and motivational variables. Though interesting, the matrix derived from Wallach and Kogan's data is obviously inadequate for any definitive study.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

THE EFFECTS OF SCHOOL AND CURRICULUM ON BOYS' ATTITUDES TOWARDS CAREERS IN ENGINEERING

By RAYMOND HUME

(*Abstract of thesis presented for the degree of M.Ed. in the University of Newcastle-upon-Tyne*, 1967)

I.—Aim.

To see if the type of secondary education which a boy receives has any influence upon his attitude towards a career in engineering.

II.—Procedure.

Three schools were chosen, each catering for the top range of abilities: an old established boys' grammar school, a boys' technical school, and a mixed grammar-technical school selected as representing a position mid-way between the other two. While the schools were different in curricula and traditions, they were as comparable as possible in size, methods of allocation, ability, range, socio-economic backgrounds of the boys, school facilities and career opportunities. A Thurstone type test of attitudes towards careers in engineering was constructed and given to a total of 102 boys in the second year and 90 boys in the fourth year of the three schools.

TABLE 1

Mean Scores on the Attitude Towards Engineering Scale.

	Grammar School	Grammar/Tech. School	Technical School
Second Years:	5·0	4·4	4·4
Fourth Years:	5·5	4·6	4·0

High scores=unfavourable attitudes.

III.—Results.

Table 1 shows changes in attitudes between the second and fourth years in each school. In the grammar and grammar-technical schools there is a movement towards more unfavourable attitudes, most marked in the grammar school. In the technical school the trend is in the opposite direction: attitudes towards careers in engineering become more favourable from the second to the fourth year. The significance of these results was examined by analysis of variance. A variance ratio of 4·1 was found between the schools as whole groups, which lies well above the 5 per cent value of 3·0.

Scores in the Morrisby test of mechanical ability for the boys in the second and fourth forms of the grammar-technical school were compared with the attitude scores. The correlations of +0·13 and +0·07 for second and fourth years respectively suggest that in the groups tested mechanical aptitude is not related to attitude towards a career in engineering.

Correlations between the attitude scale and tests from Fleming's Cotswold Personality Assessment booklet were non-significant except for the test of attitude towards working with the hands (+0·6). Correlations between attitude and school subjects proved non-significant in most cases, but there was a negative and significant correlation of −0·4 between attitude towards engineering and achievement in mathematics at the age of 13 plus.

(*Manuscript received* 19*th March*, 1968)

TEACHERS' OPINIONS ON THE PLACE OF SCIENTIFIC METHOD IN BIOLOGY COURSES

By R. AMOS

(*Abstract of thesis submitted for the degree of M.Ed., University of Manchester*, 1967.)

I.—Aim.

To ascertain the opinions of practising teachers about the objectives of biology teaching up to Ordinary Level of the General Certificate of Education, and about the difficulties experienced in teaching scientific method.

II.—Procedure.

From a 38 per cent response to a postal questionnaire sent to 674 grammar schools in the North of England, a sample of 222 teachers' replies was selected at random. There were 68 men and 68 women teachers with more than nine years' teaching experience, and 43 men and 43 women teachers with less than nine years' experience. The questionnaire required the rating of 19 objectives on five-point scales in each of three situations : actual teaching ; ideal circumstances ; and GCE Ordinary Level examinations. Teachers also rated nine difficulties met in teaching scientific method. Analyses of variance and factor analysis were applied to the responses.

III.—Results and conclusions.

The main emphasis in biology teaching up to Ordinary Level is on the imparting of information, whereas biology teachers would prefer to include more understanding of scientific method than the GCE examinations are thought to require. Although teachers adjust their teaching to the believed requirements of the GCE examinations, they pursue objectives concerned with scientific enquiry which they consider unlikely to be evaluated in the examinations. Teachers feel that the present syllabuses and external examinations, the short time allocated to biology and large classes are obstacles to the teaching of biology in a spirit of enquiry. Length of teaching experience has been shown to have considerable bearing upon the opinions expressed, and this might be connected with an increasing ability to overcome these obstacles.

(*Manuscript received* 5*th June*, 1968)

THE ROLE OF THE TEACHER OF SIX AND SEVEN-YEAR-OLD CHILDREN

By C. L. BAIRD

(*Abstract of thesis submitted for the degree of M.Ed., University of Manchester*, 1967)

I.—Aim.

To record, classify and analyse the classroom activities of a sample of 20 teachers of six and seven-year-old children, from five infant and five infant-junior schools.

II.—Procedure.

Systematic observation was carried out in each classroom during five 30-minute periods on different days in the Spring and Summer terms. The recorded data were classified by a checklist into four categories, on their contribution to: (1) control and discipline; (2) intellectual development; (3) emotional security; (4) social attitudes. Time-tables were also studied and were classified as 'progressive' or 'traditional' in emphasis. Comparisons by analysis of variance were made in relation to: (1) age-group of children; (2) infant or infant-junior school; (3) length of teaching experience; (4) progressive or traditional emphasis in time-table.

III.—Results.

Among the categories of observed behaviour, the over-all order of frequency was: (1) contribution to intellectual development; (2) contribution to emotional security; (3) fostering social attitudes; (4) maintaining control and discipline. The same category order was obtained for 90 per cent. of the sample of teachers taken individually.

The scores of teachers of seven-year-old children showed a greater proportion of behaviour contributing to intellectual development than did scores of teachers of six-year-old children. No significant difference in score was found between inexperienced and experienced teachers in infant schools; significant differences occurred between such teachers in infant-junior schools. In the infant schools, 70 per cent. of the teachers used time-tables with a progressive bias; in infant-junior schools, 80 per cent use time-tables with a traditional bias.

(*Manuscript received 5th June*, 1968)

THE ABILITIES AND INTERESTS OF CRAFT AND TECHNICIAN STUDENTS OF MECHANICAL ENGINEERING

By B. STAVELEY

(*Abstract of thesis submitted for the degree of M.Ed. of the University of Manchester,* 1967)

I.—Aim.

To compare, in technical colleges of different sizes, the proportions of students allocated initially to the Mechanical Engineering Craft Practice course and to the General Engineering course leading to the Technicians' course; to compare success rates in relation to allocation; and to discover what guidance, based on objective assessment, might lead to a more satisfactory initial allocation to the two courses.

II.—Sample.

388 students, the total intake of first-year students into part-time day Mechanical Engineering Craft Practice and General Engineering courses in each of three technical colleges in Lancashire (61, 95 and 232 students, respectively).

III.—Procedure.

The progress of each student was followed to the intermediate stage of the Technicians or the Craft Practice course. During the first term, seven tests were applied: High School Personality Questionnaire; Maudsley Personality Inventory; Devon Interest Test; General Information Test (I); Primary Mental Abilities; Arithmetic-Mathematics Test; ACER Mechanical Reasoning Test.

IV.—Results.

In the smaller colleges, a larger proportion of students was allocated to the Technicians' course, and there was a greater amount of transfer between courses. Eventual success rates were not significantly different between colleges, though there were differences in the proportions of students reaching intermediate stage in minimum time. Analysis of test scores indicated that the successful groups of students in the corresponding courses in the colleges were not significantly different in ability and that some improvement in the allocation of students to courses might be brought about by the use of objective assessment.

(*Manuscript received 5th June*, 1968)

BOOK REVIEWS

ALDERSON, C. (1968). *Magazines Teenagers Read.* London: Pergamon Press, pp. 127, 30s.

This book grew out of a special study undertaken by Mrs. Alderson when she was a student at Garnett College. Professor Bernstein who has written the Foreword amply sums up its value when he says: " Her book is an angry book with few holds barred, written by an extremely able and sensitive teacher. It does not claim to be sociological, neither does it meet the demands of the social survey. Yet I suggest that it will spark off more systematic study by those who have public funds and two or three years to think and do. Mrs. Alderson's book was written under quite other circumstances."

The early chapters study intensively the contents of three teenage magazines, *Valentine*, *Jackie* and *Trend*, as published over the period of eight months ending in March, 1966. The stories are analysed with emphasis on stock situations of plots, taboos, reader identification and anti-intellectual attitudes. Then follow chapters which deal with pop music and the pop scene, the image of the pop star, fashion, letters, advertisements. Finally, there is an interesting chapter on language which indicates how the esoteric language used helps the teenager " to believe that he belongs to an intimate club where adults and any symbols of authority are left outside."

The last part of the book consists of an analysis of the replies to 1,200 questionnaires distributed in February, 1966, to four London schools within a radius of three miles of each other. This analysis is in three parts: (*a*) comics and magazines; (*b*) newspapers; (*c*) radio and television programmes. The author is aware that the sample is obviously not representative, but some interesting pointers for further enquiry emerge.

Mrs. Alderson concludes: " We must not therefore underestimate the pernicious influence of teenage magazines, as teenagers in these magazines are treated as second-class citizens and are left ill-informed about the real world around them; they become preoccupied with entertainment and fun as a mode of life, instead of life itself. The magazines support, confirm and perpetuate the restricted lives and ideas of the teenage reader."

In spite of its technical shortcomings this book provides food for deep thought for all who have anything to do with the education and welfare of our youth.

JOHN G. LANG.

BERKOWITZ, L. (1967). *Advances in Experimental Social Psychology.* Vol. 3. New York and London: Academic Press, pp. xi+333, 84s.

This is the third volume in a series containing original papers which " try to provide a theoretical integration of findings in a particular research area." The papers are as usual of diverse orientation, though there is a clear emphasis this time on attitude structure and cognitive consistency.

Three of the six chapters deal with these topics. Janis contributes a paper on the effects of fear arousal on attitude change, the theme of which is that a simple monotonic relationship between intensity of fear arousal and change cannot account for the empirical evidence. In place of this assumption he postulates a family of inverted U-curves, and he is successful in accounting for most research data within this framework. The idea that relationships may be curvilinear is becoming an orthodox part of social psychology, and perhaps Janis could have capitalised more upon his readers' familiarity with the possibility. The two other papers dealing with communication and change each take an established model for examination and development. Feather is concerned to represent the way a person perceives and structures a communication situation in terms of the mathematical theory of linear graphs. This is not as complicated as it sounds, and an interesting attempt is made to reformulate Heider's principle of balance in linear graph terms. As the author points out, the reformulation is still in its early stages. In another chapter Tannenbaum deals with work deriving from the principle of congruity. He notes that this principle was originally developed to handle *change* of attitude, but the present chapter aims to increase understanding of how change may be *reduced*. To this end a lot of empirical matter is somewhat mystifyingly described. The account of research studies is rather difficult to follow, and it is not too clear that the congruity principle is needed to interpret the findings.

A paper by Argyle and Kendon deals with topics and studies which also are rather tenuously related to the model which supposedly integrates them. This paper is said to be one which applies to social behaviour the model of Welford and Broadbent for the analysis of sensorimotor performance. The model's relevance is not brought out very clearly, but the authors' analysis of behaviour in social situations is nevertheless an intriguing one. An account is given of work on factors influencing social performance (orientation, posture, patterning in time and visual contact, for instance). This is clearly a difficult area to research, but a most important one. The topic of Moscovici's paper may be similarly characterised. He examines the use of langauge in social interaction, and is pointedly critical of the work by psycholinguists for failing to examine the 'psychosociology' of language. (In psycholinguistic discussions, he comments, "a sort of soliloquy for several voices is substituted for true dialogue"). Moscovici himself makes a start in interpreting the kinds of language variables and the types of communication systems which are observable in social situations.

The remaining paper in the book is by Abelson. This is of a more general orientation and examines the use of mathematical models in social psychology. The author summarises and develops several mathematical techniques for sociogram analysis, interpretation of large social structures, decision processes and so on, and his accounts will doubtless be of value to people working in these areas. The chapter is however disappointing in its failure to examine the difficult logical problems of model-testing and in its remarkably abrupt conclusion. For the volume as a whole we must be grateful, but parts of its will necessarily fail to satisfy its readers.

PETER B. WARR.

COFFEY, H. S., WIENER, L. L., *et al.* (1967). *Group Treatment of Autistic Children.* London: Prentice-Hall, pp. xii + 132, 40s.

One possible answer to the social psychologist's riddle, "When is a group not a group?" would be "When it is composed of autistic children"; in which case the title of this book harbours, as its authors are aware, an internal contradiction. For how do you use essentially *group* methods with individuals who typically do not perceive, relate to, communicate with, feel about, others (or themselves) in the ways which such methods apparently take for granted? The authors sidestep this question by hoping ("formulating the hypotheses," p. 25) that if a few non-psychotic but behaviour-disordered children are introduced into the company, then the autistics will imitate social interaction from them. What happened in the group reported on here, however (as you might expect), was that these so-called 'catalysts' imitated psychotic behaviour.

Where specific technical issues are raised, we find that, with the exception of two paragraphs on the whole topic of 'group dynamics,' the therapeutic principles discussed all have a regular place in *individual* treatment also. The two paragraphs concerned (p. 47) consist mainly of remarks about the non-psychotics taking over from the sweet-distributing therapist as the group's 'focus,' and of the promise that concrete examples of group dynamics in action will be given. But when these latter are produced (pp. 63-66), most turn out to be merely descriptions of relationships between two or three individual children (such as might develop in almost any school playground), without obviously reflecting the influence or use of the group *qua* group.

The book is about a therapeutic play-group for disturbed children, organised on a day-care basis at East Bay 'Activity Center' in California. The latter is run by various mental health workers and school-teachers, but has psychiatric supervision. The domestic history of the establishment is described, and the treatment of two groups of children (about half of them psychotic) who attended it, over two years, generally a year at a time.

The authors are not concerned to put too fine a point on the concept of 'childhood autism,' and are content to indicate more-or-less ostensively, and in the light of Rimland and early Kanner, what sort of disorder they are talking about. Yet it is arguable that some aetiological attention to what sort of thing may have gone wrong developmentally would generate fairly specific suggestions about how to try to use such interactions as occur in the therapeutic situation. You may say that 'ego-boundaries' are weak or vague, but if you want to see how to strengthen or clarify them most readily it *may* be helpful to consider the origins of such weakness and vagueness. Traditional methods based on body-schema, spatial-location, positive reinforcement of interpersonal contact and so on, as used *inter alia* in this study, may be justifiable by the quality of the implied theory or by their demonstrable effectiveness. The authors almost make a virtue, however, out of their avowal that the handling of group-situations was essentially intuitive rather than theory-determined; but they do set out five guiding principles (p. 44) and advance some metaphor-laden rationale—the latter seeming to imply distinctions and conceptions which *ex hypothesi* are foreign to the psychotic predicament. In view of this, the publisher's blurb seriously misleads, to say the least, in asserting that the book "provides . . . a clearly defined examination of autism in children"; and the same applies to its reference to "the incidence of highly successful treatment of the borderline child," for there was an incidence of one.

One's patience dwindles further when a chapter pretentiously headed "The conception and philosophy of group-therapy" contains a mere four-and-a-half pages, which are separated from another chapter on virtually the same subject by one of three-and-a-half on domestic trivialities. There is a pointless photograph of one therapist "working on body image" (how, we are not really told); there is the 'famous line' of Shakespeare which gets garbled and reduced to four metrical feet; there is the demonstration that the difference between some before-and-after-treatment ratings is not statistically significant, followed by the insistence that the raw data do nevertheless suggest a 'trend' (when is a trend not a trend?); by way of references, there is a 'selected bibliography' of twelve titles; there is no index.

Therapeutic effectiveness was assessed by one-way-screen observation of, and coded commentaries on, the way children behaved throughout a session. These data are then accommodated to a rating-scale reflecting normality of interpersonal contact. They are then used to answer such questions as (i) do the children, considered (*a*) as a single group, (*b*) as separable psychotic/non-psychotic sub-groups, (*c*) as individuals, raise their average monthly rating over the treatment-period?; (ii) do the psychotics vary more than the catalysts in their month-by-month levels of 'normality'?; (iii) do the psychotics in the end relate more to peers (rather than adults) than they did to begin with? The answers are for the most part unencouraging; which is surprising, since few precautions were taken to prevent ratings being influenced by the wishes and expectations of the raters. As a result of dubiously chosen base-lines, spurious 'improvement' is displayed for two non-psychotics on criterion (i); and the figures (though not the text) indicate that one psychotic got much worse by the same token. NEIL M. CHESHIRE.

DINNAGE, R., and KELLMER PRINGLE, M. L. (1967). *Residential Child Care: Facts and Fallacies*; and *Foster Home Care: Facts and Fallacies*. London: Longmans, pp. 344 and 268, 30s. each.

These two complementary volumes continue the series of Studies in Child Development which has already given us several much needed volumes and in their general presentation are similar to the earlier issue, *Adoption: Facts and Fallacies*.

Each of these *Facts and Fallacies* books is in effect a compilation bringing together information on recent and on-going research together with a chapter of some twenty pages supplying an over-view or general survey of each particular field.

As is pointed out in their opening sections on 'Aims and Methods' these publications are part of a project sponsored by the Home Office and carried out by the National Bureau for Co-operation in Child Care. The general aim is to make available data in connection with different groups of children deprived of ordinary home life. Accordingly, the year 1948, the date of the Children Act which legislated for such children, is the point in time from which Rosemary Dinnage and M. L. Kellmer Pringle work in collecting their material.

In each of the volumes under review research concerning both the children and the adults involved in substitute care is presented. Attention to the adult sector would seem of great potential importance. Much obviously depends on the character qualities and skills that are required of child care officers, house-mothers and foster parents. But it appears from the evidence and from the commentary in the overviews that firm information is scanty. Facts in this area, as in so many other areas of child care, are difficult to establish because of the many variables involved.

Chapters in each volume list and summarise the main researches. These are clearly set out and constitute the core of each book. General and special researches are in separate sections, and descriptions of major completed studies and annotated lists of on-going work comprise different chapters. In the case of completed researches, findings are stated in some detail, though the accounts otherwise vary considerably in length, presumably in accordance with the importance of the study concerned. Little prominence is given to statistical aspects and the statistical techniques employed are rarely mentioned.

Subsequent chapters list bibliographies, and a general index is provided in each book. These indexes are more helpful than that found in *Adoption: Facts and Fallacies* but they are still rather slight.

Undoubtedly all workers in the Child Care sphere are much in the joint authors' debt for these comprehensive compilations. It is to be hoped that they will facilitate more and better research to fill in the numerous gaps in our knowledge.

O. C. SAMPSON.

FRAISSE, P., and PIAGET, J. (Eds.) (1968). *Experimental Psychology: Its Scope and Method*. Vol. V. *Motivation, Emotion and Personality*. London: Routledge and Kegan Paul, pp. 289, 35s.

This book first appeared in French in 1963, so the references, although extensive for the size and type of the book, are not entirely up to date. A doubt that may immediately be cast aside concerns the language. Each chapter is clearly and well translated although occasionally the structure and flow of sentences is French rather than English. One further important general point deserves attention: there is reference in this book to French and German work which is all too often ignored in British and American texts, as witness the neglect of Schneider and Sjobring.

The first section by Nuttin on Motivation surveys most of the work and theories in this area but sometimes refers to them so briefly that the uninformed reader (and this is manifestly a book for students) cannot really be expected to grasp the point, e.g., "He (Adler) speaks of Hortrieb, a Sehtrieb, and Ernahrungstreib, etc. However, this type of motivation should not be generalised." This chapter is too compressed and is the least satisfactory of the three.

The section on the emotions, by Fraisse, gives a very clear yet not oversimplified picture of the modern physiological view of the subject with an excellent historical survey. It is a pity that this chapter antedates some of the recent Russian work on the subject.

The section on the structure of the personality by Meili contains one of the best discussions of the factor-analytic position in personality measurement known to this reviewer. This includes a most elegant survey of the psychometric difficulties implicit in typologies and reviews not only the well known work of French, Eysenck and Cattell (embracing his work on ergs) but the studies of Heymans, Wiersma, Le Senne and Berger—which are not as widely known as they deserve to be. What is also impressive is that factor analysis is not uncritically endorsed.

All the chapters are far more than collections of relevant evidence; a clear argument and a point of view is presented in each with the empirical data on both sides discussed. All through, examples of everyday human behaviour are used to illustrate the points. The book seems, therefore, most useful as an introduction to the experimental and empirical study of these fields and the references to European work are especially valuable.

PAUL KLINE.

GLEASON, G. T. (1968). *The Theory and Nature of Independent Learning.* Pennsylvania: International Textbook Co., pp. 101, 27s.

Six papers from a conference on the theme of independent learning, and representing view-points from philosophy, psychology, sociology, anthropology, instructional technology, and phenomenology are collected in this book. Such a multi-disciplinary approach presents opportunities for cross-fertilisation, but leads also to problems of communication, and these are not overcome here. The need to establish the basics of each discipline leaves little time for addressing the topic. And in the absence of a jointly accepted definition of independent learning, the different ' varieties of discourse ' pursue parallel lines, the problem of synthesis is largely ignored, and the book remains an aggregate of individual papers.

For J. B. Macdonald, independent learning is characterised by the freedom of the learner to choose his own goals. The responsibility of the educator is to acquaint him with those viable alternatives which meet certain moral criteria—service, vitality, beauty, and justice.

R. M. Gagne mainly differentiates stimulus-response and mediational theories of learning, the latter, he feels, implying a degree of independence in all learning, since the mediational component is usually supplied by the learner. He concludes with some pertinent hypotheses relating to independent learning, regarding which, it is claimed, we have no evidence at present. A more diligent search of the literature would, one feels, have been rewarded.

In contrast to the limited perspective of Gagne's paper, Pauline Sears provides a comprehensive if unsystematic review of the relevant aspects of motivational research and theory.

In her cross-cultural approach, Dorothy Lee appears at times to confuse independent learning with learning to be independent. She indicates how some societies cope with the problem of instructing, without destroying the intellectual autonomy of the pupil. There remains some doubt whether the ' discovery ' methods she describes could be applied in a society which requires the assimilation of large amounts of culturally-defined knowledge.

As a brief review of recent technological developments in education, G. T. Gleason's paper is both systematic and comprehensive, although it contains little that is new, and less that relates exclusively to independent learning. Indeed, devices such as teaching machines, would be more accurately described as providing opportunities for individualised instruction.

S. M. Jourard, the author of the final contribution, describes his view-point as phenomenological, and he regards learning as independent when it is undertaken for oneself rather than for others. He provides a forceful if flamboyant statement of the difficulty in reconciling the confict between training and creativity, between the demands of the group and the personal development of the individual, but falls short of suggesting solutions of general applicability.

In fact, the book as a whole contains far more questions than answers, and it is difficult to recommend it wholeheartedly to anyone. The specialist will find it superficial, and to the casual reader it will appear fragmentary. It might conceivably serve as a mnemonic to someone contemplating useful research in this field.

W. B. HANTON.

Kidd, A. J., and Rivoire, J. L. (Eds.) (1966). *Perceptual Development in Children*. London: University Press, pp. 548, 100s.

This book is divided into six sections comprising physiological factors in perceptual development, theories of perception, visual and auditory perception, social, affective, and cognitive aspects of perception. Each section contains an introduction. For some sections the introduction gives a very brief summary of the contents, for others it is merely a few words. The book has a separate name and subject index.

Some of the chapters are almost equivalent to journal articles, reporting one experimental study carried out by the authors and associated research workers. The review chapters vary. Some reviews work in a very narrow area, others are specialised articles very much on the edge of research and others again are traditional, general review chapters. The editors have cast the net very wide; chapters range, for example, from one on the biochemistry at the synopses to another on the development of philosophical and moral concepts. The book will be bought and parts of it read because there is no alternative. For some time now those for whom the book is designed, 'teachers of child psychology,' have heard rumours of the important happenings in the field of perceptual development, particularly early perceptual development. In fact, there has been an increasing interest and recognition of the importance of early perceptual development but a lack of reputable work.

Certain areas are, however, relatively well developed, and these are very competently dealt with in this book. The section on visual and auditory perception is extremely valuable. It contains a full review chapter by the editors on the development of colour, space and movement in children; a chapter by one of the editors and her husband on the development of auditory perception and a chapter by Frantz on pattern discrimination and selective attention. This author develops the view that visual perception precedes action. The section on cognitive development contains review articles by M. D. Vernon and K. Lovell. These are excellent summaries of work which will be familiar to British readers.

The section on affective aspects of perception exemplifies the patchy nature of the book. Solley, in a chapter on affective processes in perceptual development, draws attention to the fact that there is more theorising than information in this area. McCarthy's article on affective aspects of language learning is a revised and extended APA address. The author puts forward an affective explanation of language development drawing attention to the fact that children's first words are exclamations, interjections, wishes, commands or threats. Her article deals with language and concept development, language and personality, strikes a note of caution concerning the influence of linguistics, discusses children with language problems and, predictably, considers language and social class albeit without a mention of Bernstein.

With regard to the development of the subject the chapter reviewing theories of perception by Pronko, Ebert and Greenberg is important. The authors review the many pre- and semi-scientific theories and criticise perceptual theorists for their pre-occupation with the 'inside.' They make a plea to get away from the age old approach and see the way ahead with something like Kantor's event-oriented approach.

The authors may be congratulated on bringing together work from such wide areas. This reviewer would have preferred a cheaper book containing general review articles on work which has been done and briefer mention of those areas where there is little experimental work.

Joan Taylor.

Komisar, B. P., and Macmillan, C. B. J. (Eds.) (1967). *Psychological Concepts in Education*. Chicago: Rand McNally, pp. 255, 38s.

" One reason why many students do not regard the canons of inquiry as binding upon them is that they have never resolved to become participants in the community of inquirers. They have a non-participating policy." This quotation from one of the sixteen analytical essays here printed or reprinted might be applied to some psychologists and educationists themselves. Some psychologists are so concerned with empirical respectability and some educationists with practical respectability that neither may relish the labour of logically analysing their own concepts in any depth. Perhaps this has been partly a reaction against one particularly uncongenial school or period of philosophy—nineteenth century idealism. Things are changing, and Komisar and Macmillan have made a valuable contribution to the change.

The whole book deserves the attention of those teaching the psychology of education, but even they—and their students more so—will find more substantial intellectual exercise in the critical perusal of some of the essays than of others. They might challenge themselves to study Robert Ennis's "A Concept of Critical Thinking" and see whether it sharpens them to the problems of practising and teaching critical thinking—a thoroughly important practical matter, to which Ennis refers at the end of his article. But anyone looking for more modest tooth-whetting exercises for student or tutor might begin with Edward Best's "The Suppressed Premiss in Educational Psychology," C. J. B. Macmillan's "The Concept of Adjustment," R. S. Peters's "The Concept of Character," two neatly opposed short essays on "Learning," by C. D. Hardie and R. W. Morshead, or J. E. McClellan on "B. F. Skinner's Philosophy of Human Nature" ("Don't worry about the youngster's motivation, just get him active"). All of these are clear and stimulating.

One could mention other contributions but perhaps a concluding reference to Edward Kelly's essay on "Transfer of Training: An Analytic Study" will suffice. Kelly's central thesis is that the concept of transfer of training serves virtually no purpose, except with reference to knowledge that may be a necessary or beneficial prerequisite to some subsequent learning. In 'transfer' experiments the control group is set to some artificial task which does not correspond to what individuals might in fact do outside a psychological laboratory. And how, in any case, does one distinguish the parts of a learning situation that affect later learning from the parts that do not? How do intent and awareness enter into the influence of earlier upon later learning? These are some of the points raised in Kelly's discussion. A study of his argument and of those on other topics throughout the book should contribute to clearer interpretations of the empirical data of educational psychology.

H. S. N. McFarland.

Miller, G. A. (1968). *The Psychology of Communication.* London: Allen Lane, The Penguin Press, pp. vii+197, 35s.

This book is one of the recently introduced series of Penguin hardbacks which includes such titles as Berger and Luckmann's *The Social Construction of Reality* and *Social Origins of Dictatorship and Democracy* by Barrington Moore. Originally published in the USA by Basic Books (1967), it brings together seven essays which sample communication problems at many different levels. Information processing, man-machine interaction, linguistic rules, and even paranormal communication are discussed in its short length. Miller's emphasis throughout is interrogative rather than didactic. We are instructed, but the facts serve as a vehicle to take us to another question. It is the merit of Miller's style that in the context of the question we can also see the silhouette of an answer. Take, for example, the last article 'Project Grammarama,' which, incidentally, is the only one that has not been published elsewhere. Miller here asks the question of why anyone should wish to conduct research into the the learning of grammatical rules underlying artificial languages. The research is described as fun—an intellectual field trip, but in the account of the games that were played by subjects and experimenter the shape of the answer appears. We learn *inter alia* that people who have discovered a rule that fits some grammatical instances in the artificial language will assume that it fits them all; that the principle of reinforcement did not seem to play any significant role; and that the opportunity to *use* the artificial language, through the introduction of a semantic component, resulted in a greater surety of knowledge on the part of the subject that he had mastered it. We are also introduced to a type of experimental material and an automated procedure whose potentiality is such that we are prepared to answer the next question, "Why would anyone undertake Project Linguarama?" by asserting that the ingenuity of Miller and his Harvard associates has provided a possible technique for the most difficult problem of all—the laboratory study of language acquisition.

Inevitably, essays drawn from such diverse sources as *Encounter*, *Psychological Review* and the *Proceedings of the National Electronics Conference* show variability in level of difficulty. The papers were prepared for very different audiences, and it is questionable whether their collation produces a book which meets the needs of any clearly defined group of readers. Professional workers in the field of psycholinguistics will already be familiar with most of the essays; readers from outside the field may find some of them over-technical and specific. The book is probably intended to provoke the thought of the general reader, and Miller's lucidity certainly helps in this direction, but how much more effective a successor to 'Language and Communication' would have been. When Miller produced that book in 1951, he was writing with a single purpose in mind, and what a superb introduction it was. After seventeen years, may we please have another?

D. J. Bruce.

Parkyn, G. W. (1967). *Success and Failure at the University, Vol. II, The Problem of Failure.* Wellington: New Zealand Council for Educational Research, pp. 234, $N.Z.3.60.

This volume contains a study of the reasons for failure of students at the four N.Z. universities. Failure rates varied according to factors such as whole or part-time, sex, and number of subjects studied, but the average among first-year full-time students was approximately 33 per cent. Parkyn's Volume I had shown correlation coefficients between general school performance and general university performance of the order of 0·54. The figures for individual subjects tended to be rather lower, with the suggestion that "The more general predictive measures are probably best for selecting students." Even the lowest 7 per cent. ability group of entrants achieved a pass rate of 41 per cent. Further selection therefore could reduce the rate of failure only if very severe. To halve the failure rate, it would have been necessary to exclude rather more than half the entrants. The great majority of students were of adequate ability (90 per cent were in the top half of the distribution of marks in School Certificate) and had been adequately prepared for beginning full-time studies at the standard currently being set by most university departments. The main reasons for failure must therefore be sought elswhere.

The book displays a methodical and open-minded approach. Each possible cause of failure is considered in turn, with research findings presented in a series of detailed statistical tables accompanied by a commentary which draws the threads of evidence together into lucid and succinct conclusions. The first broad conclusion is that for about 10 per cent. only of students was failure inevitable. About half of these had insufficient initial ability or attainments and were lucky to have passed the relevant entrance examination; for the other half their intellectual position relative to other students during the period between acquiring entrance qualifications and the end of the Stage 1 Course declined substantially so that they deservedly failed the first year examination.

For the remaining 23 per cent of students failure might well have been avoided. Parkyn suggests three main remedies:

(1) Endeavouring to increase the reliability of the examination process. More systematic use could be made of the results of shorter tests given regularly throughout the second half of the university year but, in order to increase the reliability of overall assessment, these would have to be marked independently, by persons other than the students' normal teachers. Some 8 per cent. of students fell into the category whose final marks fell below the acceptable level but whose performance during the year had fluctuated regularly above it and many of them might have been passed on review.

(2) Reviewing those students whose final marks were definitely but only slightly (less than one standard error) below the preliminary pass mark (i.e., that set prior to scrutiny of the actual examination papers), to decide whether their overall performance could nevertheless be regarded as satisfactory in that they had achieved most of the objectives of the Course. Some 7 per cent. might fall into this category. Parkyn suggests that students with poor attendance records or who had persistently made inadequate efforts might be failed but that those students who had consistently done their best might be given the benefit of the doubt. There would be advantage in trying to define more clearly the objectives which the students ought to reach.

(3) A student guidance service should be started or extended, to make early enquiries into the circumstances of students whose performance was disappointing and to consider possible remedies. The reasons for failure are almost infinitely varied but they might include :

(*a*) Course of study inappropriate to the abilities or personality of the students. (Quite a common complaint among British students, especially those pursuing a specialised scientific course, is that they have been ' educated out of the university.' By this is meant that university education has broadened the student's mind to such an extent that he can no longer endure the narrow syllabus of this specialised course.)

(*b*) Unsuitable accommodation or excessive travel between lodgings and work.

Anything up to 8 per cent might be rescued by the activities of an efficient student guidance service.

The combination of a review of the examination process and standards and a student guidance service might cut the overall failure rate from the present 33 per cent. to something more like 18 per cent.

The evidence from this volume will no doubt be studied carefully by the Committee of British Vice-Chancellors and Principals which is responsible for research into university admissions procedures. It may stimulate them to press for similar research investigations to be carried out in this country, both into the causes of failure and, more positively, into the criteria of ' success ' at universities (of which passing examinations is only one, though a most important one) and into the factors that make for success.
EDGAR ANSTEY.

SAWREY, J. M., and TELFORD, C. W. (1967). *Psychology of Adjustment.* 2nd Edit. Boston : Allyn and Bacon, Inc., pp. x+466, 57s.

This book represents a big step forward in the history of psychology as a science and is a vast improvement on the usual introductory textbooks in this subject. Instead of producing a collection of disconnected chapters, many of them irrelevant to the science of psychology, the authors of this book have organised their material around the central concept of ' adjustment,' which they consider in terms of six components or dimensions : *selective awareness*, which consists of ignoring certain aspects of the environment as well as noticing others ; *tolerance*, which is concerned with the acceptance and rejection of different aspects of the environment ; *autonomy*, which varies from rigid social conformity to compulsive ' inner direction ' ; *personal integration*, which is concerned with the coherence and organisation of feelings, beliefs, attitudes, etc. ; *impulse and behavioural control*, varying from extreme constriction of behaviour to uninhibited outward expression ; and *self-realisation*, which is concerned with the degree to which an individual realises his potential.

Under the last heading, Sawrey and Telford include the research of people like Maslow, who studied self-actualised people, and who has provided us with valuable information concerning the heights which mature man is capable of attaining. Maslow and other humanistic psychologists have introduced what has come to be known as the ' third force ' in psychology, to contrast it with the two previous major approaches—behaviourism and psychoanalysis. Hitherto, the stress has been on the study of animals and neurotics, and the resulting picture of man has hardly led to rejoicing, or to an optimistic view of his future possibilities. In the past, also, under the influence of Darwin, there has been an emphasis on adaptation to the environment, with adjustment and normality being seen in terms of social conformity. Now that healthy people have been studied, our views have been extended and broadened, and the authors of this book feel obliged to talk about the place of values in psychology, and to discuss the positive attributes of the ' good life.' All this is very relevant to the purpose of education, and should be kept in mind when discussing curricula and teaching methods.

Most of the studies cited in this book come from the fields of learning and personality, and a lengthy bibliography is appended to each of the fourteen chapters. Starting with Darwin's *The Expression of the Emotions in Man and Animals* (1872), the references span the years right up to the present, and enable the student to read around topics in greater depth.

Although the way in which the authors have organised their material provides an important advance on previous rather one-sided accounts of man, it still fails to present us with an integrated whole. More information might have been given about perception, which could well provide a unifying approach to the diverse topics discussed, but as many of the essential ingredients are available in this book, it now remains for others to reorganise them into an even more satisfactory and comprehensive unity.

G. W. H. Leytham.

WATHEN-DUNN, W. (Ed.) (1967). *Models for the Perception of Speech and Visual Form.* London: M.I.T., pp.x+470, 80s.

This is an impressive and valuable, yet in one sense a disappointing, book. It reports the proceedings of a symposium on models for the perception of speech and visual form; the aims of the symposium were to state current and new models and to test hypotheses derived from them against experimental data, and it was further hoped that " discussing models in two different perceptual areas might not only clarify and refine viable models but might also provide a cross-fertilization which could yield insights not otherwise obtainable." A number of invited papers are first presented, falling into four groups: the perception problem, visual shape perception I, speech perception, visual shape perception II. There follow six groups of contributed papers, concerned with visual recognition, experiments on speech perception, the role of experience in visual perception, perceptual organization in speech, formal models for visual perception, and statistical techniques and acoustic parameters in speech perception. From this list it is apparent that papers vary quite widely in topic, and also that the theme of some groups defies concise description. For example, what title more specific than " experiments in speech perception" could be given to a group of three papers concerned, respectively, with " voice and hearing in the bullfrog," " information and complex signal perception," and " the spectral shape in the FS–F3 region " ?

The sequence of contributed papers tends to be from empirical to theoretical, but in most groups neural, methematical and black-box models coexist with empirical data. Papers vary, too, in their emphasis on comparative and developmental aspects; and, most notably, in the knowledge presupposed in the reader. According to its jacket, the book is for " initiates to the science of perception and established workers in the field," and the novice in any one area is certainly at a disadvantage. On the other hand, some papers—for example, Gregory's—are surely too brief in their present form to be of any great use to layman or specialist, while others have content, clarity, and economy which makes them universally useful: for example, those of Mackay, Attneave, Muntz, Sternberg, Haber and Denes.

The symposium was held in November 1964, but publication of the proceedings has taken until 1967. While one cannot know how much the original papers have been edited for publication, the delay seems inexplicable, particularly since the book is produced in what appears to be photocopied typescript (with a not inconsiderable number of small typing errors) and since it comprises simply the papers themselves with virtually no report of the conference discussion. There are two papers in comment on a group; both (by Kolers and Sutherland) apply to the same nine-paper session, Kolers' paper referring to six of these papers and Sutherland's to three.

There is also a preface by Held to one group of papers; no other summaries or comments, apart from, literally, one or two points from discussion appended to specific papers. The banquet address, apparently delivered at the end of the symposium, is rightly treated as a separate paper and begins the book. According to Jacobson in that address the discussions were interesting, important, and productive. One can well believe it, but in the absence of any report the authors, from the reader's point of view, might as well not have met. In fact, reading the book is not unlike reading a large single volume of a specialist journal from cover to cover.

The editor argues that inclusion of further discussion summaries was impossible in view of the size of the proceedings. The difficulty must be admitted, but a better resolution might have been after all to deal separately with speech perception and with form perception. There are some excellent papers here, and models—a good many of them—are certainly supplied and tested against data, but the cross-fertilization, if it occurred, remains hidden. There must be more to the synthesis of two fields than their alternation in the sequence of sessions. ANN TAYLOR.

PUBLICATIONS RECEIVED

Books for review have been selected from the following. The date of publication is 1968, unless otherwise stated.

Alderson, C. *Magazines Teenagers Read.* Oxford : Pergamon, pp. 127, 30s.

Arnold, P. *Education, Physical Education and Personality Development.* London : Heineman, pp. 172, 25s.

Adcock, C. J. and N.N. *Psychology.* London : Heinemann, pp. 120, 7s. 6d.

Bigge, M., and Hunt, M. *Psychological Foundations of Education.* London : Harper and Row, pp. 603, 93s.

Blair, G., Jones, R., and Simpson, R. *Educational Psychology, 3rd Edn.* London : Collier-Macmillan, pp. 647, 65s.

Baumeister, A. (Ed.). *Mental Retardation.* London : University Press, pp. 419, 105s.

Blyth, W. A. C. *English Primary Education—Vol. 1 Schools, Vol. 2 Background.* London : Routledge Paperbacks, pp. 226 and 158, 12s. 6d. and 10s. 6d.

Bland, G. A. *Education in Special Schools for the Mentally Handicapped.* London : College of Special Education, pp. 52, 5s.

Burton, L. *Vulnerable Children.* London : Routledge and Kegan Paul, pp. 277, 35s.

Bentley, A. *Musical Education Research Papers. No. 1—Monotones.* London : Novello, pp. 74, no price.

Burkhart, R. C., and Neil, H. M. *Identity and Teacher Learning.* London : International Textbook Co., pp. 157, 45s.

Bloch, M. *Philosophie de L'Education Nouvelle.* Paris : Presses Universitaires, pp. 163, F12.

Court, G. *La Grammaire Nouvelle a L'Ecole.* Paris : Presses Universitaires, pp. 163, F12.

Cratty, B. J. *Psychology and Physical Activity.* London : Prentice-Hall, pp. 214, 55s. 6d.

Cane, B. *Teacher Education.* London: N.F.E.R., pp. 36, 5s.

Collier, K. G. *New Dimensions in Higher Education.* London : Longmans, pp. 164, 18s.

Crosby, A. *Creativity and Performance in Industrial Organisation.* London and New York : Barnes and Noble, and Tavistock, pp. 195, 42s.

Cook, J. B., *et al.* *The Search for Independence.* California : Wadsworth, pp. 104, 14s.

Dearden, R. F. *The Philosophy of Primary Education.* London : Routledge and Kegan Paul, pp. 188, 21s.

Ekman, B. *Teaching Geometry in Grade Four.* Stockholm : Scandinavian University Books, pp. 200, $3.95.

Evans, K. M. *Planning Small-Scale Research.* London : N.F.E.R., pp. 87, 6s, 6d.

Edwards, D., *General Psychology.* London : Collier-Macmillan, pp. 394, 55s.

Graven, J. *Non-Human Thought.* London : Arlington Books, pp. 202, 30s.

Gantier, H. *L'Enseignement d'une Langue Etrangere.* Paris : Presses Universitaires, pp. 142, F9.

Goodacre, E. J. *Teachers and their Pupils' Home Background.* London : N.F.E.R. pp. 170, 25s.

Gibson, J. *The Senses as Perceptual Systems.* London : Allen and Unwin, pp. 335, 40s.

Glover, E. *The Birth of the Ego.* London : Allen and Unwin, pp. 124, 20s.

Hey, R. D. and Bear, R. (Eds.). *Early Education.* Chicago : Aldine, pp. 272, $6.95.

Howells, J. G. *Theory and Practice of Family Psychiatry.* Edinburgh : Oliver and Boyd, pp. 953, £9 9s.

Hickman, A. T. *Musical Education Research Papers. No. 3—Electronic Apparatus for Music Research.* London : Novello, pp. 15, no price.

Juif, P., and Dovero, F. *Manuel Bibliographique des Sciences de L'Education.* Paris : Presses Universitaires, pp. 318, 26F.

Jungk, D. *Probleme Des Socialen Aufstiegs Berufstatiger Jugendlicher.* Stuttgart: Ferdinand Enke Verlog, pp. 268, no price.

Kind, R., and Leedham, J. *You Begin Life.* London: Longmans., pp. 31, 4s.

Kind, R., and Leedham, J. *Babies and Families.* London: Longmans, pp. 31, 4s.

Kind, R., and Leedham, J. *You Grow Up.* London: Longmans, pp. 31, 4s.

Kind, R., and Leedham, J. *Contraception.* London: Longmans, pp. 31, 4s.

Lawton, D. *Social Class, Language and Education.* London: Routledge and Kegan Paul, pp. 181, 25s.

Lindgren, H. C. *Readings in Educational Psychology.* New York: Wiley, pp. 474, 80s.

McLeish, J. *Musical **Education** Research Papers, No.* 2—*Musical Cognition.* London: Novello, pp. 10, no price.

McNally, J., and Murray, W. *Key Words to Literacy and the Teaching of Reading.* London: The Schoolmaster Publishing Co., pp. 55, 5s.

Mischel, W. *Personality and Assessment.* New York: Wiley, pp. 365, no price.

Moray, N. *An Introduction to Psychology.* London: Blackie, pp. 212, 13s. 6d.

Miller, D. *Growth to Freedom.* London: Tavistock Paperbacks, pp. 224, 15s.

Miller, G. *The Psychology of Communication.* London: Allan Lane—The Penguin Press, pp. 197, 35s.

Nuttin, J., *et al. Experimental Psychology. Vol. V—Motivation, Emotion and Personality.* London: Routledge and Kegan Paul, pp. 289, 35s.

Newson, J. and E. *Four-Years-Old in an Urban Community.* London: Allen and Unwin, pp. 569, 60s.

Naylor, G. *Psychology: An Evolutionary Introduction.* Queensland: University Press, pp. 200, $3.95.

Nooteboom, W. A. *Some Psychological Aspects of the Choreatiform Syndrome.* Assen: Vangorcum, pp, 88, f11.50.

Oakley, R. (Ed.). *New Backgrounds.* Oxford: University Press, pp. 136, 12s. 6d.

Oldfield, R. C. and Marshall, J. C. (Eds.). *Language.* London: Penguin, pp. 392, 8s. 6d.

Payne, D. *The Specification and Measurement of Learning Outcomes.* Massachusetts: Blaisdell, pp. 209, $2.75.

Piaget, J., *et al. Experimental Psychology. Vol.* 1—*History and Method.* London: Routledge and Kegan Paul, pp. 245, 35s.

Plaut, T. *Alcohol Problems.* Oxford: University Press, pp. 200, 45s. 6d.

Shields, J. B. *The Gifted Child.* London: N.F.E.R., pp. 96, 5s.

Schestakon, A. W. *L'Enseignement Programme.* Paris: Dunod, pp. 210, no price.

Seiler, T. *Die Reversibilitat in Der Entwicklung des Denkens.* Stuttgart: Klett, pp. 174, 13 d.m.

Stein, C. (1967). *Practical Psychotherapeutic Techniques.* Illinois: Charles Thompson, pp. 201, $7.50.

Thompson, R. *The Pelican History of Psychology.* London: Penguin Books, pp. 463, 7s. 6d.

Tod, R. J. N. (Ed.). *Disturbed Children.* London: Longmans, pp. 132, 12s. 6d.

Tod, R. J. N. (Ed.). *Children in Care.* London: Longmans, pp. 115, 12s. 6d.

Valentine, C. W. *The Experimental Psychology of Beauty.* London: Tavistock Paperback, pp. 438, 25s.

Wainwright, G. *Towards Efficiency in Reading.* Cambridge: University Press, pp. 101, 17s. 6d. cloth, 7s. 6d. paper.

Wason, P. C. and Johnson-Laird, P. N. *Thinking and Reasoning.* London: Penguin, pp. 428, 8s. 6d.

Welford, A. T., *et al.* (Ed.). *Society: Psychological Problems and Methods of Study.* London: Routledge Paperback, pp. 280, 15s.

Winnicott, D. W. *The Family and Individual Development.* London: Tavistock Paperback, pp. 181, 17s. 6d.

Wolman, B. *The Unconscious Mind.* London: Prentice-Hall, pp. 244, 27s. 6d.

Wright, H. *Sex and Society.* London: Allen and Unwin, pp. 140, 21s.

— *Guide Lines for Teaching, No. 2, Vocational Guidance: No. 3, Introductory Handbook on the Severely Subnormal.* London: College of Special Education, pp. 16+pp. 32, 1s. 6d. each.

INDEX TO VOLUME XXXVIII, 1968

Articles:

Anxiety and Educational Achievement: B. P. Frost .. 293
Backward and Normal Readers of Average Intelligence, Some Differences between: K. Lovell and A. Gorton .. 240
Careers among Scottish Children of High Ability, Opinions about: H. J. Butcher and H. B. Pont .. 272
Colleges of Education, Predicting Student Performance in: G. A. Cortis .. 115
Concept Attainment among the Hausa Children of Northern Nigeria, The Effect of Urbanisation upon: H. E. Poole .. 57
Concept of Equivalence, The Tristan da Cunhan Children's: A. G. Davey .. 162
Concepts in Primary School Children, The Formation and Extension of Class: E. Stones and J. R. Heslop .. 261
Conceptualisation in New Guinea, The Effect of Western Education on Science: J. R. Prince .. 64
Conservation of Illusion-Distorted Length, Operational: F. B. Murray .. 189
Creativity and Verbal Ability: A Direct Examination of their Relationship: G. P. Ginsburg and R. G. Whittemore .. 133
Deaf School Leavers, A Factorial Study of Communication and Ability in: G. W. Montgomery .. 27
Essay Grades obtained by Science and Arts Graduates in a Department of Education, A Comparison of: P. E. Richmond .. 53
Intelligence, Common Entrance and the General Certificate of Education: S. W. Hockey .. 140
Maladjustment, Symposium: Recent Research on: J. Tizard, M. Chazan, R. Davie, J. W. B. Douglas and J. M. Ross, W. Yule and M. Rutter .. 1
Motivation and School Attainment, Academic: N. J. Entwistle .. 181
Neuroticism and School Attainment—A Linear Relationship? N. J. Entwistle and S. Cunningham .. 123
Parental Attitudes and their Relation to Boys' Behaviour: H. B. Gibson .. 233
Psycholinguistic Development and the Control of Behaviour: D. Joynt and B. Cambourne .. 249
Puberty, Family Size and the Educational Progress of Girls: P. K. Poppleton .. 286
Stanford-Binet Form L-M, 3rd Revision: A Local English Study of Norms, Concurrent Validity and Social Differences: C. J. Phillips and W. J. Bannon .. 148
Teacher, The Image of the: J. H. Clarke .. 280
Teachers' Attitudes to Education: R. A. C. Oliver and H. J. Butcher .. 38
Teachers' Marks and Pupils' Expectations: A. J. Pickup and W. S. Anthony .. 302
Teaching Ability, Rater-ratee Personality in the Assessment of: K. B. Start .. 14
Teaching Approach and the Development of Divergent Thinking Abilities in Primary Schools: F. A. Haddon and H. Lytton .. 171
Test Construction—A Technique for Producing and Verifying a Normal Distribution of Items: N. France .. 21
Variations in Performance in a University Entrance Test in English: R. A. C. Oliver and D. G. Lewis .. 45

Research Notes:

Cognitive Bias in Scottish first-year Science and Arts Undergraduates: C. K. Mackay and M. B. Cameron .. 315
Conservation, The Stability of Induced: C. K. Mackay and T. M. Kilkenny .. 313
Creativity, A Note on the Wallach-Kogan Tests of: A. J. Cropley .. 197
Creativity Correlations, An Alternative to Ward's Factor Analysis of Wallach and Kogan's: F. Fee .. 319
Feminine Image Factor in Girls' Attitudes to School Subjects: F. W. Slee .. 212
Fluctuations in Children's Cognitive Test Scores over a Two-year Period: P. S. Freyburg .. 82
GCE A-level Grades with University Examinations in Chemical Engineering, The Correlation of: D. G. Bagg .. 194
Harris-Goodenough Test, A Validation of: D. J. W. Strumpfer and C. J. P. Mienie .. 96
Musical Ability, The Validity and Reliability of Bentley's Measures of: J. McLeish .. 201
Perceptual, Representational and Conceptual Characteristics of Primary School Children: D. Satterly .. 78
Referrals by School Psychologists and General Practitioners to a Child Psychiatry Clinic: P. C. Pettriciani and S. Wolff .. 89
Schema of Proportionality, The Growth of: P. Pumfrey .. 202
Secondary School Selection Procedures in Central Africa, 1962-67, A Five-Year Follow-up of: S. H. Irvine .. 208

Teacher Ratings of Extraversion and Neuroticism and Children's Inventory Responses : S. B. G. EYSENCK and A. J. PICKUP 94
Teachers, A Problem Check List for : S. LEMESHOW 310
Variations in Drawings of "A Person" and "Myself" by Hearing-Impaired and Normal Children : J. GILLIES 86
Verbal Intelligence, Visuomotor Ability and School Achievement : M. W. BRENNER and S. GILLMAN 75
WISC with Clinic Children, Validation of a Short-Form : D. W. MCKERRACHER and R. A. WATSON 205

Thesis Abstracts :
Careers in Engineering, The Effect of School and Curriculum on Boys' Attitudes Towards : R. HUME 322
Engineering, The Abilities and Interests of Craft and Technician Students of Mechanical : B. STAVELEY 324
Married Women and Young Women Students at a Day College of Education and in their First Year as Teachers : D. CASE 102
Teacher of Six and Seven-year-old Children, The Role of : C. L. BAIRD 323
Teachers' Opinions on the Place of Scientific Method in Biology Courses : R. AMOS .. 323
Training in Study Skills on the Ability to Recall and Comprehend Written Material, The Effect of : B. AZPURUA-ARRILLAGA 101
WISC and S-B Form L-M with Eight-year-old Children, Use of the : M. BRITTAIN .. 103

Authors :
AMOS, R. : Teachers' Opinions on the Place of Scientific Method in Biology Courses 323
AZPURUA-ARRILLAGA, B.: The Effect of Training in Study Skills on the Ability to Recall and Comprehend Written Material 101
BAGG, D. G. : The Correlation of G.C.E. A-level Grades with University Examinations in Chemical Engineering 194
BAIRD, C. L. : The Role of the Teacher of Six- and Seven-year-old Children 323
BRENNER, M. W., and GILLMAN, S. : Verbal Intelligence, Visuomotor Ability and School Achievement 75
BRITTAIN, M. : A Comparative Study of Use of WISC and SB with Eight-year-old Children 103
BUTCHER, H. J., and PONT, H. B. : Opinions about Carerers among Scottish Secondary School Children of High Ability 272
CASE, D. : Married Women and Young Women Students at a Day College of Education and in their First Year as Teachers 102
CHAZAN, M. : Symposium : Recent Research on Maladjustment 1
CLARKE, J. H. : The Image of the Teacher 280
CORTIS, G. A. : Predicting Student Performance in Colleges of Education 115
CROPLEY, A. J. : A Note on the Wallach-Kogan Tests of Creativity 197
DAVEY, A. G. : The Tristan da Cunhan Children's Concept of Equivalence 162
DAVIE, R. : Symposium : Recent Research on Maladjustment 1
DOUGLAS, J. W. B. and ROSS, J. M. : Symposium : Recent Research on Maladjustment 1
ENTWISTLE, N. J. : Academic Motivation and School Attainment 181
ENTWISTLE, N. J., and CUNNINGHAM, S. : Neuroticism and School Attainment—A Linear Relationship ? 123
EYSENCK, S. B. G., and PICKUP, A. J. : Teacher Ratings of Extraversion and Neuroticism and Children's Inventory Responses 94
FEE, F. : An Alternative to Ward's Factor Analysis of Wallach and Kogan's 'Creativity' Correlations (with Comment by J. Ward) 319
FRANCE, N. : Test Construction—A Technique for Producing and Verifying a Normal Distribution of items 21
FREYBERG, P. S. : Fluctuations in Children's Cognitive Test Scores over a Two-year Period 82
FROST, B. P. : Anxiety and Educational Achievement 293
GIBSON, H. B. : Parental Attitudes and their Relation to Boys' Behaviour 233
GILLIES, J. : Variations in Drawings of 'A Person' and 'Myself' by Hearing-Impaired and Normal Children 86
GINSBURG, G. P., and WHITTEMORE, R. G. : Creativity and Verbal Ability : A Direct Examination of Their Relationship 133
HADDON, F. A., and LYTTON, H. : Teaching Approach and the Development of Divergent Thinking Abilities in Primary Schools 171
HOCKEY, S. W. : Intelligence, Common Entrance and the General Certificate of Education 140

Hume, R.: The Effects of School and Curriculum on Boys' Attitudes towards Careers in Engineering 322
Irvine, S. H.: A Five-year Follow-up of Secondary School Selection Procedures in Central Africa, 1962-67 208
Joynt, D.; and Cambourne, B.: Psycholinguistic Development and the Control of Behaviour 249
Lemeshow, S.: A Problem Check List for Teachers 310
Lovell, K., and Gorton, A.: Some Differences between Backward and Normal Readers of Average Intelligence. 240
Mackay, C. K., and Cameron, Marion: Cognitive Bias in Scottish First-year Science and Arts Undergraduates 315
Mackay, C. K., and Kilkenny, T. M.: The Stability of Induced Conservation .. 313
McKerracher, D. W., and Watson, R. A.: Validation of a Short Form WISC with Clinic Children 205
McLeish, J.: The Validity and Reliability of Bentley's Measures of Musical Ability 201
Montgomery, G. W. G.: A Factorial Study of Communication and Ability in Deaf School Leavers 27
Murray, F. B.: Operational Conservation of Illusion-Distorted Length 189
Oliver, R. A. C., and Butcher, H. J.: Teachers' Attitudes to Education 38
Oliver, R. A. C., and Lewis, D. G.: Variations in Performance in a University Entrance Test in English 45
Petricciani, P. C., and Wolff, S.: Characteristics of Referrals by School Psychologists and General Practitioners to a Child Psychiatry Clinic 89
Phillips, C. J., and Bannon, W. J.: The Stanford-Binet Form L-M, 3rd Revision: A Local English Study of Norms, Concurrent Validity and Social Differences .. 148
Pickup, A. J., and Anthony, W. S.: Teachers' Marks and Pupils' Expectations .. 302
Poole, H. E.: The Effect of Urbanization upon Scientific Concept Attainment among the Hausa Children of Northern Nigeria 57
Poppleton, P. K.: Puberty, Family Size and the Educational Progress of Girls .. 286
Prince, J. R.: The Effect of Western Education on Science Conceptualisation in New Guinea 64
Pumfrey, P.: The Growth of the Schema of Proportionality 202
Satterley, D.: Perceptual, Representational and Conceptual Characteristics of Primary School Children 78
Slee, F. W.: The Feminine Image Factor in Girls' Attitudes to School Subjects .. 212
Start, K. B.: Rater-Ratee Personality in the Assessment of Teaching Ability .. 14
Staveley, B.: The Abilities and Interests of Craft and Technician Students of Mechanical Engineering 324
Stones, E., and Heslop, J. R.: The Formation and Extension of Class Concepts in Primary School Children 261
Strümpfer, D. J. W., and Mienie, C. J. P.: A Validation of the Harris-Goodenough Test 96
Tizard, J.: Symposium: Recent Research on Maladjustment 1
Yule, W., and Rutter, M.: Symposium: Recent Research on Maladjustment .. 1

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be *preceded* by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. :

I. INTRODUCTION.
II. OUTLINE OF PREVIOUS RESEARCH.
III. PLAN OF THE INVESTIGATION.
    (i) *Subjects and materials.*
    (ii) *Experimental Procedure.*
IV. STATEMENT OF RESULTS.
V. DISCUSSION AND CONCLUSIONS.
VI. REFERENCES.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.[4] When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed :

DEARBORN, W. F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development.* Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full.)

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table *and* a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

The British Journal of Social and Clinical Psychology

Volume 7, Part 3. September 1968

TAYSIR KAWWA A survey of ethnic attitudes of some British secondary school pupils

HALLA BELOFF AND SIMPSON COUPAR Some transactional perceptions of African faces

W. F. NORRISON AND D. W. CARMENT Participation and opinion-change in two person groups as related to amount of peer group support

STAN GOOCH AND M. L. KELLMER PRINGLE The association of attainment and sociometric status with reported deviant behaviour in a progressive and a traditional junior school : a pilot longitudinal study

M. G. GELDER Verbal conditioning as a measure of interpersonal influence in psychiatric interviews

K. HOPE AND T. M. CAINE The Hysteroid Obsessoid Questionnaire : a new validation

J. H. COURT AND E. GARWOLI Schizophrenic performance on a reaction-time task with increasing levels of complexity

DAISY SCHALLING AND ANNE-SOFIE ROSEN Porteus Maze differences between psychopathic and non-psychopathic criminals

G. S. TUNE A note on differences between co-operative and non-co-operative volunteer subjects

30*s. net*. ($5.50). *Annual subscription* 90*s. net* ($16.50).

CAMBRIDGE UNIVERSITY PRESS

The British Journal of Medical Psychology

Volume 41, Part 3. September 1968

DORIS M. WILLS Problems of play and mastery in the blind child

DAVID MADDISON The relevance of conjugal bereavement for preventive psychiatry

E. B. RITSON Suicide amongst alcoholics

HARRY A. WILMER The doctor-patient relationship and the issues of pity, sympathy and empathy

MICHAEL HUMPHRY Ideals and realities of the psychiatric out-patient clinic

ANDREW C. SMITH Notes on difficulties in the definition of delusion

ROBERT H. WOODY Integrating behaviour therapy and psychotherapy

WILLIAM J. DI SCIPIO Modified progressive desensitization and homosexuality

BEATRICE FOX AND WILLIAM DI SCIPIO An exploratory study in the treatment of homosexuality by combining principles from psychoanalytical theory and conditioning : theoretical and methodological considerations

VAMIK VOLKAN AND ROBERT T. CORNEY Some considerations of satellite states and satellite dreams

T. M. CAINE AND D. J. SMAIL Attitude of psychiatric patients to staff roles and treatment methods in mental hospitals

L. CHERTOK Psychiatric dialogue between East and West

PAUL KLINE Obsessional traits, obsessional symptoms and anal erotism

PAUL KLINE The validity of the Dynamic Personality Inventory

30*s. net* ($5.50). *Annual subscription* 90*s. net* ($16.50).

CAMBRIDGE UNIVERSITY PRESS

Treatment or Torture

The philosophy, techniques, and future of psychodynamics

G. SEABORN JONES

Dr. Seaborn Jones presents a stringent reappraisal of concepts and techniques in philosophy and psychology, leading to a consideration of the development of a science of psychodynamics in relation to such contiguous fields as psychiatry, psychopharmacology, ethology, and the social sciences. He explores the practical problems of establishing the discipline in university departments and of selecting and training practitioners. Material from his own experience as a conductor of analytic groups exemplifies how group methods may subserve the threefold function of therapy, education, and research. 55s net

Group Dynamics: Research and Theory

Third Edition

Edited by
DORWIN CARTWRIGHT and ALVIN ZANDER

This is a new and revised edition of one of the basic texts in social psychology. It provides an authoritative collection and summation of theory and research concerning small groups in all areas of the social structure. Since the publication of the earlier editions, research on groups has continued at an accelerated pace; new problems have been investigated, new techniques of research have been evolved, and new theoretical formulations have been proposed. The present volume reflects these changes. Fourteen of the original articles have been replaced by fourteen new ones dealing with significant recent developments, and the editors have contributed eight chapters placing the large body of empirical findings in historical and theoretical perspective. Four further chapters summarize work done on specialized problems.

'. . . certainly the most useful single book on the subject available.'
Michael Argyle, Occupational Psychology

Hardback 65s net ; Social Science Paperback 36s net

TAVISTOCK PUBLICATIONS

New... *From the Publishers of the Distinguished Mental Measurements Yearbooks*

READING TESTS AND REVIEWS

edited by OSCAR K. BUROS

Just published — a professional reference work for all reading teachers and specialists. A publishing achievement that offers a scholarly approach to choosing and using standardized reading tests . . . presented in six unusually informative sections:

THE READING TEST INDEX . . . presents a comprehensive bibliography that lists 292 reading tests, including 209 in print tests *plus* 83 out of print tests . . . 25% of the in print tests are new tests *not previously listed in any Mental Measurements Yearbook* . . . and 17 percent have been revised or supplemented since a previous MMY listing.

THE READING TEST REVIEWS . . . presents a *complete reprinting* of the reading sections of the first six MENTAL MEASUREMENTS YEARBOOKS . . . including 349 critical reviews of 158 tests by 154 reading and testing specialists . . . plus 38 excerpts from reading test reviews published in professional journals . . . and 608 references on the construction, validity and use of specific tests.

THE MMY TEST INDEX . . . presents a classified indexing of 2,802 tests listed in the six MMY's comprising a time-saving guide to 3,573 original test reviews and 358 excerpted test reviews . . . with cross references to 23,744 references on the construction, validity and use of specific tests.

THE PUBLISHERS DIRECTORY AND INDEX . . . presents current addresses of 86 publishers of in print reading tests.

THE INDEX OF TITLES . . . presents all in print reading tests *and all other tests* listed in an MMY or its companion text, *Tests In Print.*

THE INDEX OF NAMES . . . presents authors and reviewers of reading tests, authors of reading test references *plus all reviewers of tests* in an MMY.

. . . in 544
two column pages **$15.00**

Order directly from
THE GRYPHON PRESS
220 Montgomery Street • Highland Park, N. J. 08904

SPECIAL NOTE — Effective January 1, 1969 all MENTAL MEASUREMENTS YEARBOOK and TESTS IN PRINT prices will be increased to the following: 3rd MMY, $20.50; 4th MMY, $22.50; 5th MMY, $25.00; 6th MMY, $35.00; TIP, $10.00.

THE BRITISH JOURNAL OF PSYCHOLOGY

VOLUME 59, PART 4, NOVEMBER 1968

D. LEGGE AND J. M. POTTINGER Is there a motor-constancy mechanism?
H. L. WAGNER The illusions and Ganz's theory of contour displacements
B. J. FELLOWS The reverse Müller-Lyer illusion and 'enclosure'
G. H. FISHER An experimental and theoretical appraisal of the inappropriate size-depth theories of illusions
J. T. REASON Relations between motion-sickness susceptibility, the spiral after-effect and loudness estimation
D. A. ALLPORT Phenomenal simultaneity and the perceptual moment hypothesis
L. A. PEARSON and J. E. TONG Two-flash fusion threshold : the influence of age, psychophysical method, instructions, viewing conditions, sex and subject variability
J. INGLIS On the relative effects of different sources of variation in dichotic listening performance
J. S. ANTROBUS Information theory and stimulus-independent thought
N. E. WETHERICK A further study of the inferential basis of concept attainment
C. G. COSTELLO The effects of anxiety on the relation between reaction time and stimulus light intensity
C. K. MACKAY and J. MACMILLAN A comparison of stereognostic recognition in normal children and severely subnormal adults
M. H. SHELDON Deutsch's model applied to a conditional maze
K. CONNOLLY Imprinting and the following response as a function of amount of training in domestic chicks
M. DONALDSON and G. BALFOUR Less is more : a study of language comprehension in children
M. E. HEBRON A note on the predictive validity of the MPI as tested by a semantic differential technique
M. KINSBOURNE Critical notice : the analysis of higher nervous activity in man

30*s. net* ($5.50) *Annual subscription* £4 10*s. net* ($16.50)

CAMBRIDGE UNIVERSITY PRESS

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

8-24 Sampson Road North
Birmingham 11
Telephone : 021-772 5759

EDUCATIONAL
RESEARCH
IN BRITAIN

Edited by H. J. Butcher, B.A., Ph.D.
Professor of Higher Education
University of Manchester

These 18 papers by experts describe recent research in Britain in important areas of education, ranging from children's emotional development to programmed instruction and from technological education to the training of teachers. The book ends with a consolidated bibliography of more than 750 entries.

408 pages. 8½″ × 5½″ **63/- net**

UNIVERSITY OF LONDON PRESS LTD

St. Paul's House, Warwick Lane, London, E.C.4

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor W. Sluckin, Department of Psychology, The University, Leicester.

For the *British Journal of Medical Psychology*, should be sent to Dr. Thomas Freeman, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Professor J. D. Nisbet, King's College, Aberdeen.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. Audley, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. Michael Argyle (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to: Dr. Ralph Hetherington (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor Ian M. L. Hunter, Department of Psychology, University of Keele, Keele, Staffs., whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

16 DEC 1968

CONTENTS

R. Davie; J. W. B. Douglas and J. M. Ross; M. Chazan; W. Yule and M. Rutter; and J. Tizard. : Symposium (Five Articles): Recent Research on Maladjustment 1
K. B. Start : Rater-Ratee Personality in the Assessment of Teaching Ability 14
Norman France : Test Construction—A technique for producing and verifying a normal distribution of items 21
G. W. G. Montgomery : A Factorial Study of Communication and Ability in Deaf School Leavers 27
R. A. C. Oliver and H. J. Butcher : Teachers' Attitudes to Education 38
R. A. C. Oliver and D. G. Lewis : Variations in Performance in a University Entrance Test in English 45
P. E. Richmond : A Comparison of Essay Grades Obtained by Science and Arts Graduates in a Department of Education . . . 53
H. E. Poole : The Effect of Urbanization upon Scientific Concept Attainment among the Hausa Children of Northern Nigeria . 57
J. R. Prince : The Effect of Western Education on Science Conceptualisation in New Guinea 64

Research Notes:

M. W. Brenner and S. Gillman : Verbal Intelligence, Visuomotor Ability and School Achievement 75
D. Satterly : Perceptual, Representational and Conceptual Characteristics of Primary School Children 78
P. S. Freyberg : Flunctuations in Children's Cognitive Test Scores over a Two-year Period : 82
J. Gillies : Variations in Drawings of 'A Person' and 'Myself' by Hearing-Impaired and Normal Children 86
P. C. Petricciani and S. Wolff : Characteristics of Referrals by School Psychologists and General Practitioners to a Child Psychiatry Clinic . . 89
S. B. G. Eysenck and A. J. Pickup : Teacher Ratings of Extraversion and Neuroticism and Children's Inventory Responses 94
D. J. W. Strümpfer and C. J. P. Mienie : A Validation of the Harris-Goodenough Test 96

Theses Abstracts:

B. Azpurua-Arrillaga : The Effect of Training in Study Skills on the Ability to Recall and Comprehend Written Material 101
D. Case : Married Women and Young Women Students at a Day College of Education and in their First Year as Teachers. 102
M. Brittain : A Comparative Study of Use of WISC and SB with Eight-year-old Children. 103
Book Reviews 105
Publications Received 112

R ABILITATION